

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:17:40 ; Search time 27 Seconds
(without alignments)

3301.791 Million cell updates/sec

Title: US-10-069-799-5

Perfect score: 927

Sequence: 1 MSINVIKSNIQAGLNSTKS.....SSNALQPIPTQGLAPSV 927

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 14 | 1.5 | 998 | 2 I41078 | hemolysin - Escher |
| 2 | 14 | 1.5 | 998 | 2 T00227 | hemolysin A toxin |
| 3 | 14 | 1.5 | 1023 | 1 LEECA | hemolysin A - Esch |
| 4 | 14 | 1.5 | 1024 | 2 S10056 | hemolysin A - Esch |
| 5 | 14 | 1.5 | 1049 | 1 S51784 | toxin III - Actino |
| 6 | 14 | 1.5 | 1052 | 1 B49219 | toxin III - Actino |
| 7 | 11 | 1.2 | 956 | 1 B33389 | toxin II - Actinob |
| 8 | 11 | 1.2 | 956 | 1 A43834 | toxin II - Actinob |
| 9 | 11 | 1.2 | 1022 | 1 I39643 | RTX-toxin I - Acti |
| 10 | 10 | 1.1 | 953 | 1 B30169 | leukotoxin A - Pas |
| 11 | 10 | 1.1 | 955 | 1 A35254 | leukotoxin A - Pas |
| 12 | 10 | 1.1 | 1403 | 2 S77624 | mannuronan C-5-epi |
| 13 | 10 | 1.1 | 4936 | 2 A42515 | hypothetical prote |
| 14 | 8 | 0.9 | 161 | 2 AG2664 | ATP synthase B cha |
| 15 | 8 | 0.9 | 161 | 2 F97446 | hypothetical prote |
| 16 | 8 | 0.9 | 208 | 2 S34238 | leukotoxin A - Pas |
| 17 | 8 | 0.9 | 234 | 2 A00468 | conserved hypotet |
| 18 | 8 | 0.9 | 262 | 2 E70714 | hypothetical prote |
| 19 | 8 | 0.9 | 270 | 2 AC2987 | conserved hypotet |
| 20 | 8 | 0.9 | 270 | 2 E98296 | hypothetical prote |
| 21 | 8 | 0.9 | 275 | 2 J00183 | chitinase (EC 3.2. |
| 22 | 8 | 0.9 | 314 | 2 JC7335 | chitinase (EC 3.2. |
| 23 | 8 | 0.9 | 317 | 2 A03003 | hypothetical prote |
| 24 | 8 | 0.9 | 317 | 2 E98280 | mpsa protein limpo |
| 25 | 8 | 0.9 | 321 | 2 A81054 | probable membrane |
| 26 | 8 | 0.9 | 325 | 2 C87464 | lipic acid synth |
| 27 | 8 | 0.9 | 353 | 2 E95914 | probable secreted |
| 28 | 8 | 0.9 | 368 | 2 AF1434 | AA3-600 quinol oxi |
| 29 | 8 | 0.9 | 368 | 2 AF1806 | AA3-600 quinol oxi |

| | | | | | |
|-----|---|-----|------|----------|---------------------|
| 30 | 8 | 0.9 | 380 | 2 AB3160 | santhopine deamina |
| 31 | 8 | 0.9 | 394 | 2 T19181 | hypothetical prote |
| 32 | 8 | 0.9 | 399 | 2 T19180 | hypothetical prote |
| 33 | 8 | 0.9 | 433 | 1 E64242 | GTP-binding protei |
| 34 | 8 | 0.9 | 443 | 2 AE0826 | probable cadaverin |
| 35 | 8 | 0.9 | 447 | 2 C84306 | hypothetical prote |
| 36 | 8 | 0.9 | 465 | 2 H95369 | Eglic ENDO-1,3-1,4- |
| 37 | 8 | 0.9 | 469 | 2 B84372 | DNA damage-inducib |
| 38 | 8 | 0.9 | 474 | 2 S60902 | CDP-ribitol pyroph |
| 39 | 8 | 0.9 | 477 | 2 D84306 | sodium- and chlori |
| 40 | 8 | 0.9 | 505 | 2 G84079 | sodium-dependent t |
| 41 | 8 | 0.9 | 539 | 2 G95405 | hypothetical prote |
| 42 | 8 | 0.9 | 553 | 2 S77623 | mannuronan C-5-epi |
| 43 | 8 | 0.9 | 559 | 2 E84213 | hypothetical prote |
| 44 | 8 | 0.9 | 589 | 2 AD2263 | hypothetical prote |
| 45 | 8 | 0.9 | 643 | 2 T03518 | hypothetical prote |
| 46 | 8 | 0.9 | 773 | 2 S46011 | probable pre-mRNA- |
| 47 | 8 | 0.9 | 864 | 2 D70183 | chemotaxis histidi |
| 48 | 8 | 0.9 | 993 | 2 AE1905 | outer membrane sec |
| 49 | 8 | 0.9 | 997 | 2 S77625 | mannuronan C-5-epi |
| 50 | 8 | 0.9 | 997 | 2 I39739 | mannuronan C5 epim |
| 51 | 8 | 0.9 | 1055 | 1 A37205 | leukotoxin A - Act |
| 52 | 8 | 0.9 | 1112 | 2 H95964 | probable outer mem |
| 53 | 8 | 0.9 | 1208 | 2 C82779 | hemolysin-type cal |
| 54 | 8 | 0.9 | 1296 | 2 C82521 | hemolysin-type cal |
| 55 | 8 | 0.9 | 1417 | 2 AG2137 | hypothetical prote |
| 56 | 8 | 0.9 | 1705 | 2 S51672 | adenylate cyclase |
| 57 | 8 | 0.9 | 1736 | 2 F86178 | hypothetical prote |
| 58 | 8 | 0.9 | 1741 | 2 S74910 | hemolysin - Synec |
| 59 | 8 | 0.9 | 1839 | 2 S77626 | mannuronan C-5-epi |
| 60 | 8 | 0.9 | 1944 | 2 AH3098 | rhizobioicin/RTX to |
| 61 | 8 | 0.9 | 1990 | 2 A96188 | probable phosphox |
| 62 | 8 | 0.9 | 3093 | 2 AH2493 | hypothetical prote |
| 63 | 7 | 0.8 | 17 | 2 I46151 | histone H2AF - dog |
| 64 | 7 | 0.8 | 17 | 2 I45826 | histone H2AF - New |
| 65 | 7 | 0.8 | 54 | 2 S22936 | sex-determining pr |
| 66 | 7 | 0.8 | 63 | 1 CKMTA | cecropin A precurs |
| 67 | 7 | 0.8 | 63 | 2 D97714 | hypothetical prote |
| 68 | 7 | 0.8 | 79 | 2 B87640 | hypothetical prote |
| 69 | 7 | 0.8 | 79 | 2 S53337 | CD47 glycoprotein |
| 70 | 7 | 0.8 | 79 | 2 F97432 | hypothetical prote |
| 71 | 7 | 0.8 | 80 | 1 NTSR4T | neurotoxin TsIV pr |
| 72 | 7 | 0.8 | 81 | 2 G87600 | hypothetical prote |
| 73 | 7 | 0.8 | 85 | 2 E83311 | hypothetical prote |
| 74 | 7 | 0.8 | 103 | 1 G64093 | ribosomal protein |
| 75 | 7 | 0.8 | 104 | 1 R5EC24 | ribosomal protein |
| 76 | 7 | 0.8 | 104 | 2 JC2277 | ribosomal protein |
| 77 | 7 | 0.8 | 104 | 2 F91150 | 50S ribosomal subu |
| 78 | 7 | 0.8 | 104 | 2 B85996 | 50S ribosomal subu |
| 79 | 7 | 0.8 | 104 | 2 AE0027 | 50S ribosomal prot |
| 80 | 7 | 0.8 | 104 | 2 AD1007 | 50S ribosomal chai |
| 81 | 7 | 0.8 | 105 | 2 C82058 | ribosomal protein |
| 82 | 7 | 0.8 | 110 | 2 T22476 | hypothetical prote |
| 83 | 7 | 0.8 | 110 | 2 F83745 | hypothetical prote |
| 84 | 7 | 0.8 | 112 | 2 H69537 | conserved hypotet |
| 85 | 7 | 0.8 | 112 | 2 AE2244 | hypothetical prote |
| 86 | 7 | 0.8 | 118 | 2 S51207 | cruxhalorhodopsin- |
| 87 | 7 | 0.8 | 118 | 2 S19120 | pilin (pilS5 locus |
| 88 | 7 | 0.8 | 122 | 2 AD0456 | 50S ribosomal prot |
| 89 | 7 | 0.8 | 123 | 2 JE0093 | histone H2A.F/Z va |
| 90 | 7 | 0.8 | 123 | 2 H81235 | 50S ribosomal prot |
| 91 | 7 | 0.8 | 125 | 2 S07392 | histone H2A.F/Z - |
| 92 | 7 | 0.8 | 125 | 2 A28343 | sexual inducer gly |
| 93 | 7 | 0.8 | 126 | 2 G84215 | hypothetical prote |
| 94 | 7 | 0.8 | 128 | 1 HSC2HF | histone H2A.F, emb |
| 95 | 7 | 0.8 | 128 | 2 A35881 | histone H2A.Z - hu |
| 96 | 7 | 0.8 | 128 | 2 S03642 | histone H2A.Z - bo |
| 97 | 7 | 0.8 | 128 | 2 S03644 | histone H2A.Z - ra |
| 98 | 7 | 0.8 | 133 | 2 S60562 | photosystem II pro |
| 99 | 7 | 0.8 | 133 | 2 A84710 | photosystem II rea |
| 100 | 7 | 0.8 | 133 | 2 C97301 | uncharacterized co |
| 101 | 7 | 0.8 | 141 | 2 S08118 | histone H2A.VD - f |
| 102 | 7 | 0.8 | 141 | 2 S73848 | hypothetical prote |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|---|--------|---------------------|-----|---|-----|-----|---|--------|--------------------|
| 103 | 7 | 0.8 | 147 | 2 | D86389 | hypothetical prote | 176 | 7 | 0.8 | 317 | 2 | B83760 | hypothetical prote |
| 104 | 7 | 0.8 | 149 | 2 | T21037 | hypothetical prote | 177 | 7 | 0.8 | 319 | 2 | AB2121 | cysteine synthase |
| 105 | 7 | 0.8 | 151 | 2 | H90256 | hypothetical prote | 178 | 7 | 0.8 | 319 | 2 | F83402 | binding protein co |
| 106 | 7 | 0.8 | 154 | 2 | A42418 | auracyanin BI - Ch | 179 | 7 | 0.8 | 319 | 2 | S06367 | NDI intron 1 prote |
| 107 | 7 | 0.8 | 168 | 2 | S57416 | fibrillar protein M | 180 | 7 | 0.8 | 320 | 2 | S66176 | ACC oxidase (clone |
| 108 | 7 | 0.8 | 168 | 2 | A45943 | vitelline membrane | 181 | 7 | 0.8 | 320 | 2 | A40216 | flavonol 4'-sulfot |
| 109 | 7 | 0.8 | 172 | 2 | E95958 | hypothetical membr | 182 | 7 | 0.8 | 320 | 2 | AH2374 | cysteine synthase |
| 110 | 7 | 0.8 | 182 | 2 | S23604 | hypothetical prote | 183 | 7 | 0.8 | 320 | 2 | G96714 | hypothetical prote |
| 111 | 7 | 0.8 | 181 | 2 | S01150 | H+-transporting tw | 184 | 7 | 0.8 | 323 | 2 | A48997 | tumor surface anti |
| 112 | 7 | 0.8 | 186 | 2 | S35057 | NADH2 dehydrogenas | 185 | 7 | 0.8 | 324 | 2 | E83306 | cysteine synthase |
| 113 | 7 | 0.8 | 187 | 2 | S48361 | transcription fact | 186 | 7 | 0.8 | 324 | 2 | A70584 | phosphate transpor |
| 114 | 7 | 0.8 | 189 | 2 | A10213 | probable membrane | 187 | 7 | 0.8 | 326 | 2 | S54042 | hypothetical prote |
| 115 | 7 | 0.8 | 199 | 2 | A72235 | translation elonga | 188 | 7 | 0.8 | 326 | 4 | S61652 | hypothetical prote |
| 116 | 7 | 0.8 | 199 | 2 | S75344 | hypothetical prote | 189 | 7 | 0.8 | 333 | 1 | J80590 | endo-1,4-beta-xyla |
| 117 | 7 | 0.8 | 206 | 1 | I64088 | conserved hypoteth | 190 | 7 | 0.8 | 333 | 2 | I55593 | connexin 37 - huma |
| 118 | 7 | 0.8 | 208 | 2 | S01012 | sexual inducer gly | 191 | 7 | 0.8 | 334 | 2 | A29561 | prostatic spermine |
| 119 | 7 | 0.8 | 213 | 2 | B87149 | probable carbonic | 192 | 7 | 0.8 | 335 | 2 | T51106 | drpGlucose 4,6-de |
| 120 | 7 | 0.8 | 216 | 2 | C75403 | hypothetical prote | 193 | 7 | 0.8 | 338 | 2 | H87510 | polyprenyl synthet |
| 121 | 7 | 0.8 | 217 | 2 | AF3050 | two component resp | 194 | 7 | 0.8 | 338 | 2 | B82128 | conserved hypoteth |
| 122 | 7 | 0.8 | 217 | 2 | F98235 | probable two-compo | 195 | 7 | 0.8 | 338 | 2 | G71128 | hypothetical prote |
| 123 | 7 | 0.8 | 218 | 2 | S87659 | brain-specific pro | 196 | 7 | 0.8 | 339 | 2 | AC3406 | l-lactate permease |
| 124 | 7 | 0.8 | 218 | 2 | H82758 | hypothetical prote | 197 | 7 | 0.8 | 340 | 2 | A55528 | recombination prot |
| 125 | 7 | 0.8 | 219 | 2 | G83137 | riboflavin synthas | 198 | 7 | 0.8 | 341 | 2 | B86792 | hypothetical prote |
| 126 | 7 | 0.8 | 219 | 2 | A70338 | ABC transporter - | 199 | 7 | 0.8 | 348 | 2 | G83976 | cytochrome caa3 ox |
| 127 | 7 | 0.8 | 219 | 2 | B95953 | probable secreted | 200 | 7 | 0.8 | 349 | 2 | T16882 | hypothetical prote |
| 128 | 7 | 0.8 | 220 | 2 | E84057 | hypothetical prote | 201 | 7 | 0.8 | 349 | 2 | AC0419 | probable transport |
| 129 | 7 | 0.8 | 220 | 2 | F89218 | conserved hypoteth | 202 | 7 | 0.8 | 350 | 2 | AC2375 | hypothetical prote |
| 130 | 7 | 0.8 | 227 | 2 | F83097 | probable two-compo | 203 | 7 | 0.8 | 351 | 2 | B83872 | hypothetical prote |
| 131 | 7 | 0.8 | 229 | 1 | G69066 | probable 3,4-dihyd | 204 | 7 | 0.8 | 351 | 2 | D90264 | biotin synthase (b |
| 132 | 7 | 0.8 | 232 | 2 | S70355 | phosphoribosylanth | 205 | 7 | 0.8 | 351 | 2 | AC0302 | conserved hypoteth |
| 133 | 7 | 0.8 | 237 | 2 | C42606 | orfB 3' to orf405 | 206 | 7 | 0.8 | 353 | 2 | JA0123 | replicase - potato |
| 134 | 7 | 0.8 | 240 | 2 | B45359 | proopiomelanocorti | 207 | 7 | 0.8 | 355 | 2 | T11600 | NADH2 dehydrogenas |
| 135 | 7 | 0.8 | 241 | 2 | B5628 | DNA-binding stress | 208 | 7 | 0.8 | 357 | 2 | T19336 | hypothetical prote |
| 136 | 7 | 0.8 | 249 | 2 | JC2268 | hemagglutinin - Ma | 209 | 7 | 0.8 | 358 | 2 | G70601 | hypothetical prote |
| 137 | 7 | 0.8 | 247 | 2 | A13156 | short-chain dehydr | 210 | 7 | 0.8 | 360 | 2 | AD1929 | hypothetical prote |
| 138 | 7 | 0.8 | 249 | 2 | H98130 | probable short-cha | 211 | 7 | 0.8 | 362 | 2 | F97847 | recA protein [impo |
| 139 | 7 | 0.8 | 251 | 2 | AC0534 | probable hydroxyac | 212 | 7 | 0.8 | 363 | 2 | S51104 | outer membrane por |
| 140 | 7 | 0.8 | 254 | 2 | S02066 | transcription regu | 213 | 7 | 0.8 | 364 | 2 | AE2611 | conserved hypoteth |
| 141 | 7 | 0.8 | 254 | 2 | E90969 | hypothetical prote | 214 | 7 | 0.8 | 364 | 2 | C97393 | hypothetical prote |
| 142 | 7 | 0.8 | 254 | 2 | D85742 | hypothetical prote | 215 | 7 | 0.8 | 366 | 2 | B81055 | chorismate synthas |
| 143 | 7 | 0.8 | 256 | 2 | S11878 | chlorophyll a/b-bi | 216 | 7 | 0.8 | 366 | 2 | A81822 | chorismate synthas |
| 144 | 7 | 0.8 | 257 | 2 | E82094 | sodium-translocati | 217 | 7 | 0.8 | 367 | 2 | C82391 | N-ethylmaleimide r |
| 145 | 7 | 0.8 | 258 | 1 | CTONPK | corticotropin / li | 218 | 7 | 0.8 | 367 | 2 | T12790 | N-acetyluramoyl-L |
| 146 | 7 | 0.8 | 260 | 1 | ORXL3 | cytochrome-c oxida | 219 | 7 | 0.8 | 371 | 2 | T51695 | cell division prot |
| 147 | 7 | 0.8 | 261 | 2 | T11540 | cytochrome-c oxida | 220 | 7 | 0.8 | 374 | 2 | E82890 | chorismate synthas |
| 148 | 7 | 0.8 | 261 | 2 | T11460 | cytochrome-c oxida | 221 | 7 | 0.8 | 375 | 2 | E75542 | probable endogluc |
| 149 | 7 | 0.8 | 261 | 2 | T11826 | cytochrome-c oxida | 222 | 7 | 0.8 | 375 | 2 | JQ0846 | DNA-binding protei |
| 150 | 7 | 0.8 | 262 | 2 | T34691 | hypothetical prote | 223 | 7 | 0.8 | 376 | 2 | G73941 | oligopeptide trans |
| 151 | 7 | 0.8 | 269 | 2 | D70660 | probable arOE - My | 224 | 7 | 0.8 | 379 | 2 | G75428 | conserved hypoteth |
| 152 | 7 | 0.8 | 275 | 2 | A22826 | phytohemagglutinin | 225 | 7 | 0.8 | 380 | 2 | T42755 | tyrosylprotein sul |
| 153 | 7 | 0.8 | 276 | 2 | S51831 | phytohemagglutinin | 226 | 7 | 0.8 | 382 | 2 | E64686 | myosin-like protei |
| 154 | 7 | 0.8 | 276 | 2 | T48843 | halorhodopsin [val | 227 | 7 | 0.8 | 384 | 2 | T38544 | probable exopolyp |
| 155 | 7 | 0.8 | 278 | 2 | E83453 | hypothetical prote | 228 | 7 | 0.8 | 385 | 2 | D95350 | probable transmem |
| 156 | 7 | 0.8 | 279 | 2 | B71415 | probable phosphoch | 229 | 7 | 0.8 | 386 | 2 | D71851 | tetracycline resis |
| 157 | 7 | 0.8 | 279 | 2 | B87434 | hypothetical prote | 230 | 7 | 0.8 | 386 | 2 | E64665 | tetracycline resis |
| 158 | 7 | 0.8 | 281 | 2 | AG1369 | conserved hypoteth | 231 | 7 | 0.8 | 387 | 2 | E95933 | probable calcium-b |
| 159 | 7 | 0.8 | 282 | 2 | T25037 | hypothetical prote | 232 | 7 | 0.8 | 388 | 2 | AC3216 | isomerase/lactoniz |
| 160 | 7 | 0.8 | 284 | 2 | S08385 | nodO protein - Rhi | 233 | 7 | 0.8 | 388 | 2 | AG2723 | phage-related inte |
| 161 | 7 | 0.8 | 284 | 2 | A43721 | nodule formation p | 234 | 7 | 0.8 | 389 | 2 | A60109 | major outer membra |
| 162 | 7 | 0.8 | 291 | 2 | AB1025 | arac family regula | 235 | 7 | 0.8 | 393 | 2 | AF0365 | probable aminotran |
| 163 | 7 | 0.8 | 291 | 2 | H72238 | methicillin resist | 236 | 7 | 0.8 | 393 | 2 | H90441 | hypothetical prote |
| 164 | 7 | 0.8 | 294 | 2 | AG0913 | probable adenine-s | 237 | 7 | 0.8 | 397 | 2 | F83385 | probable MFS trans |
| 165 | 7 | 0.8 | 296 | 2 | T26406 | hypothetical prote | 238 | 7 | 0.8 | 400 | 2 | C82855 | aromatic-amino-aci |
| 166 | 7 | 0.8 | 297 | 2 | S75656 | hypothetical prote | 239 | 7 | 0.8 | 401 | 2 | B84060 | multidrug resistan |
| 167 | 7 | 0.8 | 298 | 2 | E85166 | probable phosphoch | 240 | 7 | 0.8 | 404 | 1 | G69318 | threonine synthase |
| 168 | 7 | 0.8 | 299 | 2 | S76631 | hypothetical prote | 241 | 7 | 0.8 | 406 | 1 | JNEBPT | phosphoglycerate t |
| 169 | 7 | 0.8 | 302 | 2 | D82351 | radD protein VC019 | 242 | 7 | 0.8 | 406 | 2 | C43664 | tryptophan synthas |
| 170 | 7 | 0.8 | 307 | 2 | S73632 | Holliday junction | 243 | 7 | 0.8 | 407 | 2 | B70185 | probable hydroxyme |
| 171 | 7 | 0.8 | 307 | 2 | G84239 | Holliday junction | 244 | 7 | 0.8 | 407 | 2 | D86782 | peptidoglycan bios |
| 172 | 7 | 0.8 | 311 | 2 | G96596 | hypothetical prote | 245 | 7 | 0.8 | 410 | 2 | F87688 | tryptophan synthas |
| 173 | 7 | 0.8 | 311 | 2 | D72693 | hypothetical prote | 246 | 7 | 0.8 | 410 | 2 | G85991 | hypothetical prote |
| 174 | 7 | 0.8 | 312 | 2 | T16001 | hypothetical prote | 247 | 7 | 0.8 | 411 | 2 | H95851 | probable alcohol d |
| 175 | 7 | 0.8 | 314 | 2 | F96527 | protein F27J15.20 | 248 | 7 | 0.8 | 414 | 1 | JQ1016 | glycine hydroxymet |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|---|--------|---------------------------------|-----|---|-----|------|---|--------|--|
| 249 | 7 | 0.8 | 414 | 2 | H81383 | glycine hydroxymet | 322 | 7 | 0.8 | 585 | 1 | F64159 | hypothetical prote |
| 250 | 7 | 0.8 | 414 | 2 | E83489 | hypothetical prote | 323 | 7 | 0.8 | 587 | 2 | E82155 | sigma-54 dependent |
| 251 | 7 | 0.8 | 415 | 2 | A12087 | hypothetical prote | 324 | 7 | 0.8 | 589 | 2 | T25895 | hypothetical prote |
| 252 | 7 | 0.8 | 426 | 2 | E87385 | chromate transport | 325 | 7 | 0.8 | 589 | 2 | B97806 | conserved hypothet |
| 253 | 7 | 0.8 | 429 | 2 | C97023 | homoserine dehydro | 326 | 7 | 0.8 | 594 | 2 | C81011 | probable periplasm |
| 254 | 7 | 0.8 | 430 | 2 | C97237 | histidyl-tRNA synt | 327 | 7 | 0.8 | 594 | 2 | H81954 | hypothetical prote |
| 255 | 7 | 0.8 | 431 | 2 | G86277 | F14L17.11 protein | 328 | 7 | 0.8 | 595 | 2 | T26843 | two-component sens |
| 256 | 7 | 0.8 | 439 | 2 | T01270 | hypothetical prote | 329 | 7 | 0.8 | 600 | 2 | B83875 | prSD protein (U891 |
| 257 | 7 | 0.8 | 441 | 2 | G75306 | drug transport pro | 330 | 7 | 0.8 | 606 | 2 | H98187 | zinc metalloprotei |
| 258 | 7 | 0.8 | 442 | 2 | H70960 | hypothetical prote | 331 | 7 | 0.8 | 608 | 2 | B82938 | probable Arp /Gfp |
| 259 | 7 | 0.8 | 443 | 2 | AF1901 | hypothetical prote | 332 | 7 | 0.8 | 609 | 2 | B81385 | hypothetical prote |
| 260 | 7 | 0.8 | 444 | 2 | E83033 | probable MFS trans | 333 | 7 | 0.8 | 616 | 2 | AG2957 | probable c4-dicarb |
| 261 | 7 | 0.8 | 449 | 2 | F72383 | dihydrolipoamide d | 334 | 7 | 0.8 | 616 | 2 | G98325 | hypothetical prote |
| 262 | 7 | 0.8 | 449 | 2 | G64349 | phosphomannomutase | 335 | 7 | 0.8 | 617 | 2 | C75407 | excinuclease ABC c |
| 263 | 7 | 0.8 | 449 | 2 | B86783 | hypothetical prote | 336 | 7 | 0.8 | 620 | 2 | F72395 | hypothetical prote |
| 264 | 7 | 0.8 | 450 | 2 | C86262 | F13K23.12 protein | 337 | 7 | 0.8 | 632 | 2 | AG1366 | phosphotransferase |
| 265 | 7 | 0.8 | 452 | 2 | H90249 | pyruvate kinase (p | 338 | 7 | 0.8 | 632 | 2 | AH1735 | phosphotransferase |
| 266 | 7 | 0.8 | 452 | 2 | B59096 | hypothetical prote | 339 | 7 | 0.8 | 635 | 2 | F90551 | lipoprotein [impor |
| 267 | 7 | 0.8 | 452 | 2 | F75217 | hypothetical prote | 340 | 7 | 0.8 | 637 | 2 | T50951 | hypothetical prote |
| 268 | 7 | 0.8 | 454 | 2 | E75116 | 4-aminobutyrate am | 341 | 7 | 0.8 | 639 | 2 | A97450 | 1-deoxy-D-xylulose |
| 269 | 7 | 0.8 | 454 | 2 | A71016 | probable 4-aminobu | 342 | 7 | 0.8 | 639 | 2 | AC2668 | 1-deoxy-D-xylulose |
| 270 | 7 | 0.8 | 455 | 2 | G71860 | biotin carboxylase | 343 | 7 | 0.8 | 643 | 2 | AD3439 | 1-deoxyxylulose-5- |
| 271 | 7 | 0.8 | 458 | 2 | B64566 | biotin carboxylase | 344 | 7 | 0.8 | 646 | 2 | S01351 | type III site-spec |
| 272 | 7 | 0.8 | 460 | 2 | B83552 | probable two-compo | 345 | 7 | 0.8 | 646 | 2 | T23039 | hypothetical prote |
| 273 | 7 | 0.8 | 460 | 2 | I57546 | Rabin3 - rat | 346 | 7 | 0.8 | 654 | 2 | H86717 | hypothetical prote |
| 274 | 7 | 0.8 | 462 | 2 | A83893 | two-component sens | 347 | 7 | 0.8 | 655 | 2 | T34705 | hypothetical prote |
| 275 | 7 | 0.8 | 463 | 2 | AB0807 | phosphoglycerate t | 348 | 7 | 0.8 | 658 | 2 | H81676 | hypothetical prote |
| 276 | 7 | 0.8 | 463 | 2 | T51313 | hypothetical prote | 349 | 7 | 0.8 | 658 | 2 | D96656 | hypothetical prote |
| 277 | 7 | 0.8 | 465 | 2 | AB2995 | endo-1,3-1,4-beta- | 350 | 7 | 0.8 | 660 | 2 | A40158 | dnak-type molecula |
| 278 | 7 | 0.8 | 466 | 2 | T45880 | hypothetical prote | 351 | 7 | 0.8 | 660 | 2 | B71521 | dnak-type molecula |
| 279 | 7 | 0.8 | 467 | 2 | D84938 | H ₂ -transporting tw | 352 | 7 | 0.8 | 662 | 2 | T32821 | hypothetical prote |
| 280 | 7 | 0.8 | 471 | 2 | T34956 | probable UDP-N-ace | 353 | 7 | 0.8 | 670 | 2 | A25001 | catalase (EC 1.11. |
| 281 | 7 | 0.8 | 473 | 2 | G90401 | flagella-related p | 354 | 7 | 0.8 | 671 | 2 | E82355 | ATP-dependent DNA |
| 282 | 7 | 0.8 | 473 | 2 | AE3584 | exopolysphatase | 355 | 7 | 0.8 | 676 | 2 | A12058 | hypothetical prote |
| 283 | 7 | 0.8 | 474 | 2 | AB0002 | probable membrane | 356 | 7 | 0.8 | 685 | 2 | G97351 | K ⁺ -transporting Ar |
| 284 | 7 | 0.8 | 475 | 2 | S48132 | metalloproteinase | 357 | 7 | 0.8 | 685 | 2 | T46843 | H ⁺ /K ⁺ -exchanging A |
| 285 | 7 | 0.8 | 477 | 2 | D82587 | cationic amino aci | 358 | 7 | 0.8 | 689 | 2 | AC1408 | transcription anti |
| 286 | 7 | 0.8 | 477 | 4 | I41144 | aspartate ammonia- | 359 | 7 | 0.8 | 689 | 2 | AC1784 | transcription anti |
| 287 | 7 | 0.8 | 478 | 1 | UFEC0W | aspartate ammonia- | 360 | 7 | 0.8 | 698 | 2 | S49206 | Gl cyclin CLN1 - y |
| 288 | 7 | 0.8 | 479 | 2 | A38307 | metalloproteinase | 361 | 7 | 0.8 | 703 | 2 | A49994 | sulfate transporte |
| 289 | 7 | 0.8 | 482 | 2 | T08256 | cytochrome d ubiqu | 362 | 7 | 0.8 | 711 | 2 | A90023 | DNA topoisomerase |
| 290 | 7 | 0.8 | 482 | 2 | H98288 | endo-1,3-1,4-beta- | 363 | 7 | 0.8 | 726 | 2 | B72665 | hypothetical prote |
| 291 | 7 | 0.8 | 483 | 2 | S37055 | catalase (EC 1.11. | 364 | 7 | 0.8 | 735 | 2 | S76425 | hypothetical prote |
| 292 | 7 | 0.8 | 483 | 2 | G87412 | outer membrane pro | 365 | 7 | 0.8 | 745 | 2 | G01025 | serine/threonine p |
| 293 | 7 | 0.8 | 484 | 2 | C82171 | cardiolipin syntha | 366 | 7 | 0.8 | 746 | 2 | B70359 | transcription regu |
| 294 | 7 | 0.8 | 485 | 2 | T25199 | hypothetical prote | 367 | 7 | 0.8 | 756 | 2 | C87432 | hypothetical prote |
| 295 | 7 | 0.8 | 486 | 1 | HYSE15 | serralysin (EC 3.4 | 368 | 7 | 0.8 | 757 | 2 | F83516 | hypothetical prote |
| 296 | 7 | 0.8 | 488 | 1 | H64313 | corrinoid/iron-sul | 369 | 7 | 0.8 | 764 | 2 | T05409 | hypothetical prote |
| 297 | 7 | 0.8 | 492 | 2 | A55589 | lysine-tRNA ligase | 370 | 7 | 0.8 | 767 | 2 | C97763 | coat protein Vp1 - |
| 298 | 7 | 0.8 | 493 | 2 | H91288 | aspartate ammonia- | 371 | 7 | 0.8 | 781 | 1 | VCPV19 | neurofilament trip |
| 299 | 7 | 0.8 | 493 | 2 | F86109 | aspartate ammonia- | 372 | 7 | 0.8 | 784 | 2 | PN0009 | hypothetical prote |
| 300 | 7 | 0.8 | 500 | 2 | B83910 | hypothetical prote | 373 | 7 | 0.8 | 790 | 2 | AE2203 | hypothetical prote |
| 301 | 7 | 0.8 | 504 | 2 | S12164 | serralysin (EC 3.4 | 374 | 7 | 0.8 | 792 | 2 | S03232 | probable [imported |
| 302 | 7 | 0.8 | 505 | 2 | H85361 | leucyl aminopeptid | 375 | 7 | 0.8 | 806 | 2 | G95362 | hypothetical prote |
| 303 | 7 | 0.8 | 505 | 2 | T04143 | CLB1 protein - tom | 376 | 7 | 0.8 | 808 | 2 | H64474 | hypothetical prote |
| 304 | 7 | 0.8 | 506 | 2 | T22876 | hypothetical prote | 377 | 7 | 0.8 | 815 | 2 | H69775 | hypothetical prote |
| 305 | 7 | 0.8 | 510 | 2 | T50526 | CaLB protein - Ara | 378 | 7 | 0.8 | 817 | 1 | TGHUM1 | protein-glutamine |
| 306 | 7 | 0.8 | 512 | 2 | H70707 | probable ggTA prot | 379 | 7 | 0.8 | 827 | 1 | S10639 | fructose phosphotr |
| 307 | 7 | 0.8 | 525 | 2 | A87468 | malate synthase [i | 380 | 7 | 0.8 | 843 | 2 | T34618 | NADH2 dehydrogenas |
| 308 | 7 | 0.8 | 530 | 2 | T21171 | hypothetical prote | 381 | 7 | 0.8 | 847 | 2 | T19544 | hypothetical prote |
| 309 | 7 | 0.8 | 532 | 1 | QFPGM | neurofilament trip | 382 | 7 | 0.8 | 850 | 2 | G70332 | conserved hypothet |
| 310 | 7 | 0.8 | 534 | 2 | F72682 | hypothetical prote | 383 | 7 | 0.8 | 855 | 2 | A71223 | hypothetical prote |
| 311 | 7 | 0.8 | 547 | 2 | A48442 | membrane transport | 384 | 7 | 0.8 | 855 | 2 | B75191 | probable ATP-depen |
| 312 | 7 | 0.8 | 557 | 1 | S34412 | transcription fact | 385 | 7 | 0.8 | 855 | 2 | D96594 | unknown protein, 7 |
| 313 | 7 | 0.8 | 557 | 1 | S15342 | transcription fact | 386 | 7 | 0.8 | 899 | 2 | D96594 | hypothetical prote |
| 314 | 7 | 0.8 | 557 | 2 | T47128 | heat shock protein | 387 | 7 | 0.8 | 900 | 2 | T41607 | conserved hypothet |
| 315 | 7 | 0.8 | 558 | 1 | A39633 | transcription fact | 388 | 7 | 0.8 | 902 | 2 | AC1852 | hypothetical prote |
| 316 | 7 | 0.8 | 558 | 2 | E72220 | conserved hypothet | 389 | 7 | 0.8 | 906 | 2 | T08723 | hypothetical prote |
| 317 | 7 | 0.8 | 559 | 2 | S33724 | transcription fact | 390 | 7 | 0.8 | 922 | 2 | S38723 | glutamate receptor |
| 318 | 7 | 0.8 | 559 | 2 | D95913 | probable secreted | 391 | 7 | 0.8 | 933 | 2 | C55226 | receptor-like prot |
| 319 | 7 | 0.8 | 575 | 2 | T48224 | probable homeodoma | 392 | 7 | 0.8 | 1014 | 2 | H81894 | cylM protein - Ent |
| 320 | 7 | 0.8 | 576 | 2 | A70900 | hypothetical glyci | 393 | 7 | 0.8 | 1014 | 2 | H81125 | probable cell-divi |
| 321 | 7 | 0.8 | 583 | 2 | A13098 | hypothetical prote | 394 | 7 | 0.8 | 1030 | 2 | S57380 | probable membrane |
| | | | | | | | | | | 1034 | 2 | D65119 | acriflavin resista |

| | | | | | | | | | | | | |
|-----|---|-----|------|---|--------|-----|---|-----|----|---|--------|--------------------|
| 395 | 7 | 0.8 | 1038 | 2 | JC5497 | 468 | 6 | 0.6 | 65 | 2 | S19568 | parsin, ovary-matu |
| 396 | 7 | 0.8 | 1072 | 2 | G35851 | 469 | 6 | 0.6 | 65 | 2 | G72172 | H8R protein - vari |
| 397 | 7 | 0.8 | 1083 | 2 | A38919 | 470 | 6 | 0.6 | 66 | 2 | D82805 | hypothetical prote |
| 398 | 7 | 0.8 | 1109 | 2 | A56143 | 471 | 6 | 0.6 | 67 | 2 | B70214 | hypothetical prote |
| 399 | 7 | 0.8 | 1131 | 2 | D75429 | 472 | 6 | 0.6 | 68 | 2 | D85716 | probable holin pro |
| 400 | 7 | 0.8 | 1138 | 2 | A48944 | 473 | 6 | 0.6 | 68 | 2 | D90766 | probable holin pro |
| 401 | 7 | 0.8 | 1146 | 2 | S59376 | 474 | 6 | 0.6 | 68 | 2 | T17853 | hypothetical prote |
| 402 | 7 | 0.8 | 1170 | 2 | A33595 | 475 | 6 | 0.6 | 69 | 2 | B46238 | hypothetical prote |
| 403 | 7 | 0.8 | 1202 | 2 | S55553 | 476 | 6 | 0.6 | 70 | 2 | C95284 | hypothetical prote |
| 404 | 7 | 0.8 | 1204 | 2 | T18812 | 477 | 6 | 0.6 | 71 | 1 | H70799 | integrase-related |
| 405 | 7 | 0.8 | 1208 | 2 | T42574 | 478 | 6 | 0.6 | 71 | 2 | B90357 | hypothetical prote |
| 406 | 7 | 0.8 | 1209 | 1 | DNBECA | 479 | 6 | 0.6 | 71 | 2 | AG1985 | hypothetical prote |
| 407 | 7 | 0.8 | 1237 | 2 | E86457 | 480 | 6 | 0.6 | 74 | 2 | H90657 | hypothetical prote |
| 408 | 7 | 0.8 | 1287 | 2 | B53739 | 481 | 6 | 0.6 | 74 | 2 | H85508 | hypothetical prote |
| 409 | 7 | 0.8 | 1288 | 2 | E71884 | 482 | 6 | 0.6 | 74 | 2 | AD0231 | hypothetical prote |
| 410 | 7 | 0.8 | 1289 | 2 | E90098 | 483 | 6 | 0.6 | 75 | 2 | AH1550 | hypothetical prote |
| 411 | 7 | 0.8 | 1290 | 2 | G64630 | 484 | 6 | 0.6 | 75 | 2 | A12226 | hypothetical prote |
| 412 | 7 | 0.8 | 1291 | 1 | A48940 | 485 | 6 | 0.6 | 76 | 1 | C69331 | hypothetical prote |
| 413 | 7 | 0.8 | 1291 | 2 | S44983 | 486 | 6 | 0.6 | 76 | 2 | T36168 | hypothetical prote |
| 414 | 7 | 0.8 | 1305 | 2 | D82923 | 487 | 6 | 0.6 | 76 | 2 | D83916 | hypothetical prote |
| 415 | 7 | 0.8 | 1428 | 2 | AC2224 | 488 | 6 | 0.6 | 76 | 2 | D82787 | hypothetical prote |
| 416 | 7 | 0.8 | 1525 | 2 | T14961 | 489 | 6 | 0.6 | 78 | 2 | E85682 | unknown protein en |
| 417 | 7 | 0.8 | 1609 | 2 | E87243 | 490 | 6 | 0.6 | 78 | 2 | AB1162 | flagellar switch p |
| 418 | 7 | 0.8 | 1632 | 2 | C70752 | 491 | 6 | 0.6 | 78 | 2 | AB1521 | weakly flagellar s |
| 419 | 7 | 0.8 | 1636 | 2 | B92736 | 492 | 6 | 0.6 | 79 | 2 | A86517 | hypothetical prote |
| 420 | 7 | 0.8 | 1774 | 2 | T17421 | 493 | 6 | 0.6 | 79 | 2 | F83880 | acyl-carrier prote |
| 421 | 7 | 0.8 | 1830 | 2 | E82909 | 494 | 6 | 0.6 | 79 | 2 | B72106 | hypothetical prote |
| 422 | 7 | 0.8 | 2064 | 2 | G82562 | 495 | 6 | 0.6 | 79 | 2 | A83944 | hypothetical prote |
| 423 | 7 | 0.8 | 2171 | 2 | E86347 | 496 | 6 | 0.6 | 80 | 2 | C49050 | T-cell surface gly |
| 424 | 7 | 0.8 | 2364 | 2 | A56577 | 497 | 6 | 0.6 | 80 | 2 | T08876 | phosphoribosylpro |
| 425 | 7 | 0.8 | 2427 | 2 | T16613 | 498 | 6 | 0.6 | 80 | 2 | AB3302 | hypothetical prote |
| 426 | 7 | 0.8 | 2464 | 1 | QRMSP1 | 499 | 6 | 0.6 | 81 | 1 | JN0462 | adenoregulin precu |
| 427 | 7 | 0.8 | 2761 | 2 | T21064 | 500 | 6 | 0.6 | 81 | 2 | A75384 | hypothetical prote |
| 428 | 7 | 0.8 | 2894 | 2 | C64474 | 501 | 6 | 0.6 | 81 | 2 | D81565 | hypothetical prote |
| 429 | 7 | 0.8 | 3135 | 2 | A48584 | 502 | 6 | 0.6 | 83 | 2 | A64463 | hypothetical prote |
| 430 | 7 | 0.8 | 3430 | 1 | GNWVKV | 503 | 6 | 0.6 | 83 | 2 | A45271 | invasin secretion |
| 431 | 7 | 0.8 | 3433 | 1 | GNWVKV | 504 | 6 | 0.6 | 83 | 2 | AI0303 | hypothetical prote |
| 432 | 7 | 0.8 | 3436 | 2 | S55659 | 505 | 6 | 0.6 | 84 | 1 | A26873 | small acid-soluble |
| 433 | 7 | 0.8 | 3678 | 2 | S28916 | 506 | 6 | 0.6 | 84 | 1 | Q1BFS2 | uracil-DNA glycosy |
| 434 | 6 | 0.6 | 12 | 2 | A29169 | 507 | 6 | 0.6 | 84 | 2 | A64172 | exodeoxyribonuclea |
| 435 | 6 | 0.6 | 17 | 2 | A29834 | 508 | 6 | 0.6 | 84 | 2 | F90941 | hypothetical prote |
| 436 | 6 | 0.6 | 23 | 2 | B54226 | 509 | 6 | 0.6 | 84 | 2 | B85790 | hypothetical prote |
| 437 | 6 | 0.6 | 26 | 2 | G37396 | 510 | 6 | 0.6 | 84 | 2 | A64940 | hypothetical prote |
| 438 | 6 | 0.6 | 26 | 4 | I59183 | 511 | 6 | 0.6 | 85 | 2 | G83302 | conserved hypotet |
| 439 | 6 | 0.6 | 31 | 2 | A58793 | 512 | 6 | 0.6 | 85 | 2 | AD2229 | hypothetical prote |
| 440 | 6 | 0.6 | 31 | 2 | A34448 | 513 | 6 | 0.6 | 86 | 2 | T17859 | hypothetical prote |
| 441 | 6 | 0.6 | 31 | 2 | S65418 | 514 | 6 | 0.6 | 86 | 2 | S73048 | hypothetical prote |
| 442 | 6 | 0.6 | 35 | 2 | D97553 | 515 | 6 | 0.6 | 86 | 2 | S20471 | class V zygote-spe |
| 443 | 6 | 0.6 | 36 | 2 | C94175 | 516 | 6 | 0.6 | 87 | 2 | AH1828 | hypothetical prote |
| 444 | 6 | 0.6 | 37 | 2 | S11202 | 517 | 6 | 0.6 | 88 | 2 | A82920 | hypothetical prote |
| 445 | 6 | 0.6 | 38 | 2 | S07140 | 518 | 6 | 0.6 | 89 | 2 | D72581 | hypothetical prote |
| 446 | 6 | 0.6 | 42 | 2 | S21027 | 519 | 6 | 0.6 | 92 | 2 | T36228 | hypothetical prote |
| 447 | 6 | 0.6 | 45 | 2 | A38666 | 520 | 6 | 0.6 | 93 | 2 | T01876 | hypothetical prote |
| 448 | 6 | 0.6 | 51 | 2 | S71030 | 521 | 6 | 0.6 | 94 | 2 | S24795 | ubiquinol-cytochro |
| 449 | 6 | 0.6 | 52 | 2 | B90673 | 522 | 6 | 0.6 | 94 | 2 | T06417 | pathogenesis-relat |
| 450 | 6 | 0.6 | 52 | 2 | F85523 | 523 | 6 | 0.6 | 94 | 2 | S72613 | chaperonin groES - |
| 451 | 6 | 0.6 | 54 | 2 | AC0896 | 524 | 6 | 0.6 | 94 | 2 | T30377 | hypothetical prote |
| 452 | 6 | 0.6 | 55 | 2 | D97080 | 525 | 6 | 0.6 | 94 | 2 | T03285 | anther-specific pr |
| 453 | 6 | 0.6 | 57 | 2 | F97843 | 526 | 6 | 0.6 | 94 | 2 | F95917 | hypothetical prote |
| 454 | 6 | 0.6 | 58 | 2 | AC0327 | 527 | 6 | 0.6 | 95 | 2 | B69422 | hypothetical prote |
| 455 | 6 | 0.6 | 59 | 2 | S72821 | 528 | 6 | 0.6 | 95 | 2 | T03697 | SAR8.2a protein, T |
| 456 | 6 | 0.6 | 59 | 2 | D97219 | 529 | 6 | 0.6 | 95 | 2 | T03699 | SAR8.2c protein, T |
| 457 | 6 | 0.6 | 59 | 2 | D98151 | 530 | 6 | 0.6 | 95 | 2 | T02969 | hypothetical prote |
| 458 | 6 | 0.6 | 60 | 1 | LBRFAB | 531 | 6 | 0.6 | 95 | 2 | T02969 | hypothetical prote |
| 459 | 6 | 0.6 | 60 | 2 | F90769 | 532 | 6 | 0.6 | 95 | 2 | A83880 | acyl-carrier prote |
| 460 | 6 | 0.6 | 60 | 2 | A64006 | 533 | 6 | 0.6 | 95 | 2 | A11024 | probable GerS fami |
| 461 | 6 | 0.6 | 60 | 2 | F97702 | 534 | 6 | 0.6 | 95 | 2 | C98255 | hypothetical prote |
| 462 | 6 | 0.6 | 61 | 2 | F95014 | 535 | 6 | 0.6 | 96 | 2 | G81019 | chaperonin, 10 kDa |
| 463 | 6 | 0.6 | 61 | 2 | H57887 | 536 | 6 | 0.6 | 96 | 2 | A89891 | conserved hypotet |
| 464 | 6 | 0.6 | 64 | 2 | H81811 | 537 | 6 | 0.6 | 96 | 2 | C69225 | conserved hypotet |
| 465 | 6 | 0.6 | 64 | 2 | H97099 | 538 | 6 | 0.6 | 96 | 2 | B83815 | hypothetical prote |
| 466 | 6 | 0.6 | 65 | 1 | B25234 | 539 | 6 | 0.6 | 96 | 2 | B82321 | conserved hypotet |
| 467 | 6 | 0.6 | 65 | 2 | T28606 | 540 | 6 | 0.6 | 96 | 2 | S10069 | regulatory protein |

| | | | | | | | | | | |
|-----|---|-----|-----|---|--------|--------------------|-----|---|--------|--------------------|
| 541 | 6 | 0.6 | 96 | 2 | A64626 | hypotheical prote | 614 | 2 | H90263 | transcriptional re |
| 542 | 6 | 0.6 | 97 | 2 | E87330 | S4 domain protein | 615 | 2 | T03701 | SAR8_2e protein, T |
| 543 | 6 | 0.6 | 98 | 1 | T17351 | NADH2 dehydrogenas | 616 | 2 | AB2722 | nitrogen host f |
| 544 | 6 | 0.6 | 98 | 2 | T17136 | NADH2 dehydrogenas | 617 | 2 | B69468 | protein hde8 precu |
| 545 | 6 | 0.6 | 98 | 2 | T17143 | NADH2 dehydrogenas | 618 | 2 | S30269 | hypotheical prote |
| 546 | 6 | 0.6 | 98 | 2 | T17174 | NADH2 dehydrogenas | 619 | 2 | E91177 | hypotheical prote |
| 547 | 6 | 0.6 | 98 | 2 | T17171 | NADH2 dehydrogenas | 620 | 2 | F86023 | unknown protein en |
| 548 | 6 | 0.6 | 98 | 2 | T17177 | NADH2 dehydrogenas | 621 | 2 | E85686 | hypotheical prote |
| 549 | 6 | 0.6 | 98 | 2 | T17183 | NADH2 dehydrogenas | 622 | 2 | T30494 | hypotheical prote |
| 550 | 6 | 0.6 | 98 | 2 | T17348 | NADH2 dehydrogenas | 623 | 2 | S76866 | hypotheical prote |
| 551 | 6 | 0.6 | 98 | 2 | T17360 | NADH2 dehydrogenas | 624 | 2 | T36789 | hypotheical prote |
| 552 | 6 | 0.6 | 98 | 2 | T17363 | NADH2 dehydrogenas | 625 | 1 | RIIDS1 | somatostatin-14 pr |
| 553 | 6 | 0.6 | 98 | 2 | T17147 | NADH2 dehydrogenas | 626 | 6 | I50798 | preprosomatostatin |
| 554 | 6 | 0.6 | 98 | 2 | T17154 | NADH2 dehydrogenas | 627 | 6 | C72759 | hypotheical prote |
| 555 | 6 | 0.6 | 98 | 2 | S62433 | zinc finger protei | 628 | 2 | A23329 | ly-5-8 glycoprotei |
| 556 | 6 | 0.6 | 98 | 2 | C97749 | hypotheical prote | 629 | 6 | S06809 | protein secretion |
| 557 | 6 | 0.6 | 99 | 2 | T37184 | hypotheical prote | 630 | 6 | AE3313 | hypotheical prote |
| 558 | 6 | 0.6 | 99 | 2 | T44977 | oligopeptide ABC t | 631 | 2 | T30042 | hypotheical prote |
| 559 | 6 | 0.6 | 100 | 2 | T30673 | hypotheical prote | 632 | 6 | R5YM18 | ribosomal protein |
| 560 | 6 | 0.6 | 100 | 2 | C64575 | hypotheical prote | 633 | 6 | G97503 | hypotheical prote |
| 561 | 6 | 0.6 | 101 | 2 | E82849 | conserved hypotet | 634 | 6 | E72551 | hypotheical prote |
| 562 | 6 | 0.6 | 101 | 2 | T45325 | hypotheical prote | 635 | 6 | D81945 | probable pilus bio |
| 563 | 6 | 0.6 | 101 | 2 | D97386 | hypotheical prote | 636 | 6 | F85957 | unknown protein en |
| 564 | 6 | 0.6 | 101 | 2 | AD2604 | conserved hypotet | 637 | 6 | E85605 | hypotheical prote |
| 565 | 6 | 0.6 | 101 | 2 | S04714 | hypotheical prote | 638 | 6 | B90796 | hypotheical prote |
| 566 | 6 | 0.6 | 101 | 2 | S30493 | Sp1 protein - mous | 639 | 6 | D91112 | hypotheical prote |
| 567 | 6 | 0.6 | 101 | 2 | B96904 | hypotheical prote | 640 | 6 | H90190 | probable transcrip |
| 568 | 6 | 0.6 | 101 | 2 | A97625 | hypotheical prote | 641 | 6 | T08906 | hypotheical prote |
| 569 | 6 | 0.6 | 102 | 1 | QPTK | acylphosphatase (E | 642 | 6 | D81442 | 50S ribosomal prot |
| 570 | 6 | 0.6 | 102 | 1 | QPTK | acylphosphatase (E | 643 | 6 | S19123 | hypotheical prote |
| 571 | 6 | 0.6 | 102 | 2 | AF3395 | NADH2 dehydrogenas | 644 | 6 | B71456 | hypotheical prote |
| 572 | 6 | 0.6 | 102 | 2 | A45675 | acylphosphatase (E | 645 | 6 | T42313 | hypotheical prote |
| 573 | 6 | 0.6 | 102 | 2 | C87282 | conserved hypotet | 646 | 6 | T12498 | hypotheical prote |
| 574 | 6 | 0.6 | 102 | 2 | T25332 | hypotheical prote | 647 | 6 | C70382 | ribosomal protein |
| 575 | 6 | 0.6 | 102 | 2 | T36095 | hypotheical prote | 648 | 6 | S01253 | hypotheical prote |
| 576 | 6 | 0.6 | 103 | 1 | R5BS2B | ribosomal protein | 649 | 6 | I64028 | hypotheical prote |
| 577 | 6 | 0.6 | 103 | 2 | T11995 | LSU ribosomal prot | 650 | 6 | F75110 | hypotheical prote |
| 578 | 6 | 0.6 | 103 | 2 | AB3348 | LSU ribosomal prot | 651 | 6 | B75331 | conserved hypotet |
| 579 | 6 | 0.6 | 103 | 2 | S81081 | M protein precurs | 652 | 6 | D75324 | hypotheical prote |
| 580 | 6 | 0.6 | 103 | 2 | C72307 | conserved hypotet | 653 | 6 | R5EC7 | ribosomal protein |
| 581 | 6 | 0.6 | 103 | 2 | S26067 | hypotheical prote | 654 | 6 | R5EB12 | ribosomal protein |
| 582 | 6 | 0.6 | 103 | 2 | B47048 | plasmid stabilizat | 655 | 6 | E91242 | 50S ribosomal subu |
| 583 | 6 | 0.6 | 103 | 2 | G86911 | hypotheical prote | 656 | 6 | E82336 | ribosomal protein |
| 584 | 6 | 0.6 | 103 | 2 | H90234 | hypotheical prote | 657 | 6 | B82946 | ribosomal protein |
| 585 | 6 | 0.6 | 103 | 2 | C26424 | stylar glycoprotei | 658 | 6 | B86090 | 50S ribosomal subu |
| 586 | 6 | 0.6 | 104 | 2 | S53942 | probable membrane | 659 | 6 | AF0933 | 50S ribosomal chal |
| 587 | 6 | 0.6 | 104 | 2 | E69805 | hypotheical prote | 660 | 6 | E71345 | conserved hypotet |
| 588 | 6 | 0.6 | 104 | 2 | H90411 | hypotheical prote | 661 | 6 | R5HG12 | ribosomal protein |
| 589 | 6 | 0.6 | 105 | 1 | S40199 | ribosomal protein | 662 | 6 | A24753 | ribosomal protein |
| 590 | 6 | 0.6 | 105 | 2 | G95911 | probable ferredoxi | 663 | 6 | F82533 | 50S ribosomal prot |
| 591 | 6 | 0.6 | 105 | 2 | F72678 | hypotheical prote | 664 | 6 | S49638 | hypotheical prote |
| 592 | 6 | 0.6 | 105 | 2 | A82985 | hypotheical prote | 665 | 6 | D87555 | hypotheical prote |
| 593 | 6 | 0.6 | 105 | 2 | T37176 | hypotheical prote | 666 | 6 | B97081 | iron-dependent tra |
| 594 | 6 | 0.6 | 105 | 2 | G72705 | hypotheical prote | 667 | 6 | VHWHE | structural protein |
| 595 | 6 | 0.6 | 106 | 2 | C64111 | probable translati | 668 | 6 | C44212 | structural protein |
| 596 | 6 | 0.6 | 106 | 2 | H72547 | probable periplasm | 669 | 6 | F64083 | ribosomal protein |
| 597 | 6 | 0.6 | 106 | 2 | D72652 | hypotheical prote | 670 | 6 | G97286 | ribosomal protein |
| 598 | 6 | 0.6 | 106 | 2 | T2969 | hypotheical 11.9K | 671 | 6 | G83980 | transcription regu |
| 599 | 6 | 0.6 | 106 | 2 | S72815 | hypotheical prote | 672 | 6 | D90488 | conserved hypotet |
| 600 | 6 | 0.6 | 106 | 2 | T06685 | hypotheical prote | 673 | 6 | S75382 | hypotheical prote |
| 601 | 6 | 0.6 | 106 | 2 | PN0468 | hypotheical prote | 674 | 6 | F82716 | hypotheical prote |
| 602 | 6 | 0.6 | 107 | 2 | S75164 | hypotheical prote | 675 | 6 | I40348 | ribosomal protein |
| 603 | 6 | 0.6 | 107 | 2 | U02035 | hypotheical 11.7K | 676 | 6 | I40350 | ribosomal protein |
| 604 | 6 | 0.6 | 108 | 2 | E75335 | hypotheical prote | 677 | 6 | AF3345 | LSU ribosomal prot |
| 605 | 6 | 0.6 | 108 | 2 | E64962 | hypotheical prote | 678 | 6 | B70472 | flagellar protein |
| 606 | 6 | 0.6 | 109 | 2 | S61255 | genome polyprote | 679 | 6 | B96960 | holo-acyl-carrier |
| 607 | 6 | 0.6 | 109 | 2 | T02762 | anther specific pr | 680 | 6 | T13559 | hypotheical prote |
| 608 | 6 | 0.6 | 110 | 2 | PH0211 | probable transamin | 681 | 6 | CCAGA6 | cytochrome c556 - |
| 609 | 6 | 0.6 | 110 | 2 | D87016 | phosphoribosyl-ATP | 682 | 6 | H71723 | ribosomal protein |
| 610 | 6 | 0.6 | 110 | 2 | D83170 | hypotheical prote | 683 | 6 | AC2817 | 50S ribosomal prot |
| 611 | 6 | 0.6 | 110 | 2 | E72423 | hypotheical prote | 684 | 6 | D97595 | ribosomal protein |
| 612 | 6 | 0.6 | 110 | 2 | E72663 | hypotheical prote | 685 | 6 | B97022 | general secretion |
| 613 | 6 | 0.6 | 111 | 2 | T06418 | pathogenesis-relat | 686 | 6 | S23214 | hypotheical prote |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|---|--------|--------------------|-----|---|-----|-----|---|--------|----------------------|
| 687 | 6 | 0.6 | 126 | 1 | RSDV7 | ribosomal protein | 760 | 6 | 0.6 | 141 | 1 | HADKAW | hemoglobin alpha-A |
| 688 | 6 | 0.6 | 126 | 2 | F69787 | hypothetical prote | 761 | 6 | 0.6 | 141 | 1 | HAGS | hemoglobin alpha-A |
| 689 | 6 | 0.6 | 126 | 2 | D69293 | conserved hypotet | 762 | 6 | 0.6 | 141 | 1 | HAGSC | hemoglobin alpha-A |
| 690 | 6 | 0.6 | 127 | 2 | S53854 | ribosomal protein | 763 | 6 | 0.6 | 141 | 1 | HAWS | hemoglobin alpha-A |
| 691 | 6 | 0.6 | 127 | 2 | H87310 | ribosomal protein | 764 | 6 | 0.6 | 141 | 1 | HACGAA | hemoglobin alpha-A |
| 692 | 6 | 0.6 | 127 | 2 | E86158 | 60S ribosomal prot | 765 | 6 | 0.6 | 141 | 1 | HACGAA | hemoglobin alpha-A |
| 693 | 6 | 0.6 | 127 | 2 | T47929 | hypothetical prote | 766 | 6 | 0.6 | 141 | 1 | HAGLAB | hemoglobin alpha-A |
| 694 | 6 | 0.6 | 127 | 2 | AG2667 | conserved hypotet | 767 | 6 | 0.6 | 141 | 1 | HADLA | hemoglobin alpha-A |
| 695 | 6 | 0.6 | 127 | 2 | D97449 | hypothetical prote | 768 | 6 | 0.6 | 141 | 1 | HADLA | hemoglobin alpha-A |
| 696 | 6 | 0.6 | 127 | 2 | PC2036 | microfibril-associ | 769 | 6 | 0.6 | 141 | 1 | HAHTAB | hemoglobin alpha-A |
| 697 | 6 | 0.6 | 127 | 2 | AG1425 | hypothetical secre | 770 | 6 | 0.6 | 141 | 1 | HAHTAB | hemoglobin alpha-A |
| 698 | 6 | 0.6 | 128 | 2 | D71906 | diacylglycerol kin | 771 | 6 | 0.6 | 141 | 1 | HADKDD | hemoglobin alpha-D |
| 699 | 6 | 0.6 | 128 | 2 | D64607 | diacylglycerol kin | 772 | 6 | 0.6 | 141 | 1 | HADKDD | hemoglobin alpha-D |
| 700 | 6 | 0.6 | 128 | 2 | S77671 | streptokinase A (E | 773 | 6 | 0.6 | 141 | 1 | HAGSDA | hemoglobin alpha-D |
| 701 | 6 | 0.6 | 128 | 2 | S77679 | streptokinase A (E | 774 | 6 | 0.6 | 141 | 1 | HAGSDA | hemoglobin alpha-D |
| 702 | 6 | 0.6 | 128 | 2 | S77688 | streptokinase A (E | 775 | 6 | 0.6 | 141 | 1 | HAGSDI | hemoglobin alpha-D |
| 703 | 6 | 0.6 | 128 | 2 | AE3399 | integration host f | 776 | 6 | 0.6 | 141 | 1 | HAGSDC | hemoglobin alpha-D |
| 704 | 6 | 0.6 | 128 | 2 | E70466 | ribosomal protein | 777 | 6 | 0.6 | 141 | 1 | A24625 | hemoglobin alpha-A |
| 705 | 6 | 0.6 | 128 | 2 | S13069 | ribosomal protein | 778 | 6 | 0.6 | 141 | 2 | A40463 | integrin alpha-6 c |
| 706 | 6 | 0.6 | 128 | 2 | UJ1002 | keratin, claw - ch | 779 | 6 | 0.6 | 141 | 2 | C69018 | methyl viologen-re |
| 707 | 6 | 0.6 | 128 | 2 | AC0111 | glycine cleavage s | 780 | 6 | 0.6 | 141 | 2 | A30315 | oxalacetate decar |
| 708 | 6 | 0.6 | 129 | 2 | S72292 | ribosomal protein | 781 | 6 | 0.6 | 141 | 2 | A44464 | hypothetical prote |
| 709 | 6 | 0.6 | 129 | 2 | D64444 | hypothetical prote | 782 | 6 | 0.6 | 141 | 2 | C70572 | phage-related prot |
| 710 | 6 | 0.6 | 129 | 2 | H82735 | hypothetical prote | 783 | 6 | 0.6 | 141 | 2 | F69948 | hemoglobin alpha-1 |
| 711 | 6 | 0.6 | 129 | 2 | F90349 | conserved hypotet | 784 | 6 | 0.6 | 141 | 2 | HART1 | hemoglobin alpha-1 |
| 712 | 6 | 0.6 | 129 | 2 | T20172 | hypothetical prote | 785 | 6 | 0.6 | 141 | 2 | HADK | hemoglobin alpha-A |
| 713 | 6 | 0.6 | 130 | 2 | AG2468 | 50S ribosomal prot | 786 | 6 | 0.6 | 141 | 2 | HADKAY | hemoglobin alpha-A |
| 714 | 6 | 0.6 | 130 | 2 | AB1511 | hypothetical prote | 787 | 6 | 0.6 | 141 | 2 | S39908 | S-receptor kinase |
| 715 | 6 | 0.6 | 130 | 2 | H84943 | hypothetical prote | 788 | 6 | 0.6 | 141 | 2 | S13452 | hemoglobin alpha-A |
| 716 | 6 | 0.6 | 130 | 2 | D84507 | hypothetical prote | 789 | 6 | 0.6 | 141 | 2 | A44777 | profilin spCoell - |
| 717 | 6 | 0.6 | 130 | 2 | AC1799 | hypothetical secre | 790 | 6 | 0.6 | 141 | 2 | B65049 | hypothetical prote |
| 718 | 6 | 0.6 | 131 | 2 | T07390 | 14-3-3 protein tft | 791 | 6 | 0.6 | 141 | 2 | G91072 | hypothetical prote |
| 719 | 6 | 0.6 | 131 | 2 | H97780 | hypothetical prote | 792 | 6 | 0.6 | 141 | 2 | B85917 | hypothetical prote |
| 720 | 6 | 0.6 | 131 | 2 | F90153 | hypothetical prote | 793 | 6 | 0.6 | 141 | 2 | AC1642 | hypothetical prote |
| 721 | 6 | 0.6 | 131 | 2 | D82222 | hypothetical prote | 794 | 6 | 0.6 | 141 | 2 | G87647 | hypothetical prote |
| 722 | 6 | 0.6 | 132 | 2 | T18337 | icMD protein - Leg | 795 | 6 | 0.6 | 141 | 2 | F72475 | hypothetical prote |
| 723 | 6 | 0.6 | 132 | 2 | S69753 | hypothetical prote | 796 | 6 | 0.6 | 141 | 2 | WZB332 | gene 32 protein - |
| 724 | 6 | 0.6 | 132 | 2 | C93298 | hypothetical prote | 797 | 6 | 0.6 | 141 | 2 | S47832 | hypothetical prote |
| 725 | 6 | 0.6 | 132 | 2 | T23596 | hypothetical prote | 798 | 6 | 0.6 | 141 | 2 | A98190 | hypothetical prote |
| 726 | 6 | 0.6 | 133 | 2 | GPBPP4 | gop protein - eate | 799 | 6 | 0.6 | 141 | 2 | B86037 | hypothetical prote |
| 727 | 6 | 0.6 | 133 | 2 | D87690 | hypothetical prote | 800 | 6 | 0.6 | 141 | 2 | AG0374 | probable secreted |
| 728 | 6 | 0.6 | 133 | 2 | A70986 | hypothetical prote | 801 | 6 | 0.6 | 141 | 2 | D97017 | ferric uptake regu |
| 729 | 6 | 0.6 | 134 | 2 | B90284 | hemoglobin alpha-2 | 802 | 6 | 0.6 | 141 | 2 | S66915 | hypothetical prote |
| 730 | 6 | 0.6 | 134 | 2 | S66694 | histone H2A.H7A3 - | 803 | 6 | 0.6 | 141 | 2 | D81707 | hypothetical prote |
| 731 | 6 | 0.6 | 134 | 2 | B71359 | conserved hypotet | 804 | 6 | 0.6 | 141 | 2 | H70424 | pilin - Aquifex ae |
| 732 | 6 | 0.6 | 134 | 2 | F64015 | hypothetical prote | 805 | 6 | 0.6 | 141 | 2 | T23735 | hypothetical prote |
| 733 | 6 | 0.6 | 134 | 2 | S77129 | hypothetical prote | 806 | 6 | 0.6 | 141 | 2 | F95116 | GtrA family protei |
| 734 | 6 | 0.6 | 134 | 2 | B83804 | small multidrug ex | 807 | 6 | 0.6 | 141 | 2 | D97986 | mesH proteinase (imp |
| 735 | 6 | 0.6 | 135 | 2 | AF0865 | conserved hypotet | 808 | 6 | 0.6 | 141 | 2 | JN0480 | phospholipase A2 (|
| 736 | 6 | 0.6 | 135 | 2 | T28703 | hypothetical prote | 809 | 6 | 0.6 | 141 | 2 | AF2292 | hypothetical prote |
| 737 | 6 | 0.6 | 135 | 2 | H87410 | hypothetical prote | 810 | 6 | 0.6 | 141 | 2 | PSBOA | phospholipase A2 (|
| 738 | 6 | 0.6 | 135 | 2 | I64147 | lactoylglutathione | 811 | 6 | 0.6 | 141 | 2 | TPQJ1 | troponin T, embryo |
| 739 | 6 | 0.6 | 135 | 2 | D72612 | hypothetical prote | 812 | 6 | 0.6 | 141 | 2 | TPQJ2 | troponin T, embryo |
| 740 | 6 | 0.6 | 136 | 2 | T10795 | acyl carrier prote | 813 | 6 | 0.6 | 141 | 2 | E70168 | flagellar protein |
| 741 | 6 | 0.6 | 136 | 2 | C81436 | biopolymer transpo | 814 | 6 | 0.6 | 141 | 2 | D82189 | hypothetical prote |
| 742 | 6 | 0.6 | 136 | 2 | AS5083 | hypothetical prote | 815 | 6 | 0.6 | 141 | 2 | D84407 | iron-dependent rep |
| 743 | 6 | 0.6 | 136 | 2 | E37963 | hypothetical prote | 816 | 6 | 0.6 | 141 | 2 | C34588 | DNA-directed RNA p |
| 744 | 6 | 0.6 | 137 | 2 | G86789 | hypothetical prote | 817 | 6 | 0.6 | 141 | 2 | D65134 | hypothetical prote |
| 745 | 6 | 0.6 | 137 | 2 | AC1817 | hypothetical prote | 818 | 6 | 0.6 | 141 | 2 | C91158 | hypothetical prote |
| 746 | 6 | 0.6 | 137 | 2 | C87332 | hypothetical prote | 819 | 6 | 0.6 | 141 | 2 | A86004 | hypothetical prote |
| 747 | 6 | 0.6 | 137 | 2 | AD1827 | hypothetical prote | 820 | 6 | 0.6 | 141 | 2 | B64950 | copper homeostasis |
| 748 | 6 | 0.6 | 138 | 2 | AB1306 | transcription regu | 821 | 6 | 0.6 | 141 | 2 | H90951 | copper homeostasis |
| 749 | 6 | 0.6 | 139 | 1 | H84005 | conserved hypotet | 822 | 6 | 0.6 | 141 | 2 | D85800 | copper homeostasis |
| 750 | 6 | 0.6 | 140 | 1 | HBFGRE | hemoglobin beta ch | 823 | 6 | 0.6 | 141 | 2 | AI0733 | probable bacteriop |
| 751 | 6 | 0.6 | 140 | 2 | PH0134 | Ig lambda chain pr | 824 | 6 | 0.6 | 141 | 2 | F69026 | conserved hypotet |
| 752 | 6 | 0.6 | 140 | 2 | D97722 | 50S ribosomal prot | 825 | 6 | 0.6 | 141 | 2 | F82735 | hypothetical prote |
| 753 | 6 | 0.6 | 140 | 2 | F70515 | probable lipoprote | 826 | 6 | 0.6 | 141 | 2 | E82523 | hypothetical prote |
| 754 | 6 | 0.6 | 140 | 2 | E71871 | DNA transfer prote | 827 | 6 | 0.6 | 141 | 2 | S73443 | hypothetical prote |
| 755 | 6 | 0.6 | 140 | 2 | C97204 | probable acetylra | 828 | 6 | 0.6 | 141 | 2 | T23839 | hypothetical prote |
| 756 | 6 | 0.6 | 140 | 2 | AB1678 | transcription regu | 829 | 6 | 0.6 | 141 | 2 | B69813 | hypothetical prote |
| 757 | 6 | 0.6 | 141 | 1 | HAPN | hemoglobin alpha-A | 830 | 6 | 0.6 | 141 | 2 | E69032 | formate hydrogenly |
| 758 | 6 | 0.6 | 141 | 1 | HAGSM | hemoglobin alpha-A | 831 | 6 | 0.6 | 141 | 2 | T39949 | probable copper tr |
| 759 | 6 | 0.6 | 141 | 1 | HADKAM | hemoglobin alpha-A | 832 | 6 | 0.6 | 141 | 2 | S39910 | S-receptor kinase |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|---|--------|--------------------|-----|---|-----|-----|---|--------|---------------------|
| 833 | 6 | 0.6 | 149 | 2 | S31391 | interleukin-2 prec | 906 | 6 | 0.6 | 162 | 2 | C90952 | hypothetical prote |
| 834 | 6 | 0.6 | 149 | 2 | B40463 | integrin alpha-6 c | 907 | 6 | 0.6 | 162 | 2 | G85800 | hypothetical prote |
| 835 | 6 | 0.6 | 149 | 2 | T37099 | hypothetical prote | 908 | 6 | 0.6 | 162 | 2 | S75277 | hypothetical prote |
| 836 | 6 | 0.6 | 149 | 2 | T70599 | hypothetical prote | 909 | 6 | 0.6 | 162 | 2 | G90958 | probable tail prot |
| 837 | 6 | 0.6 | 149 | 2 | S73351 | adhesin p1 homolog | 910 | 6 | 0.6 | 162 | 2 | AB1950 | hypothetical prote |
| 838 | 6 | 0.6 | 149 | 2 | B86475 | hypothetical prote | 911 | 6 | 0.6 | 163 | 1 | RDLBD | dihydrofolate redu |
| 839 | 6 | 0.6 | 149 | 2 | S22209 | photosystem I chai | 912 | 6 | 0.6 | 163 | 2 | T12561 | glycine cleavage s |
| 840 | 6 | 0.6 | 149 | 2 | T05934 | probable jasmonate | 913 | 6 | 0.6 | 163 | 2 | C71547 | hypothetical prote |
| 841 | 6 | 0.6 | 150 | 2 | S31411 | S-receptor kinase- | 914 | 6 | 0.6 | 163 | 2 | S58818 | hypothetical prote |
| 842 | 6 | 0.6 | 150 | 2 | F65100 | hypothetical 16.3 | 915 | 6 | 0.6 | 164 | 2 | A85807 | probable tail fibe |
| 843 | 6 | 0.6 | 150 | 2 | A91128 | hypothetical prote | 916 | 6 | 0.6 | 164 | 2 | A89780 | conserved hypotcet |
| 844 | 6 | 0.6 | 150 | 2 | H85972 | hypothetical prote | 917 | 6 | 0.6 | 165 | 1 | XUBSM1 | methylated-DNA- lpr |
| 845 | 6 | 0.6 | 150 | 2 | S58172 | mithramycin polyke | 918 | 6 | 0.6 | 165 | 2 | C44938 | cysteine proteinas |
| 846 | 6 | 0.6 | 150 | 2 | S78334 | photosystem I chai | 919 | 6 | 0.6 | 165 | 2 | AF2696 | hypothetical prote |
| 847 | 6 | 0.6 | 150 | 2 | T47315 | hypothetical prote | 920 | 6 | 0.6 | 165 | 2 | F97478 | ccME (AR176798) li |
| 848 | 6 | 0.6 | 151 | 2 | F89973 | conserved hypotcet | 921 | 6 | 0.6 | 165 | 2 | C84460 | hypothetical prote |
| 849 | 6 | 0.6 | 152 | 2 | S00332 | alpha-amylase inhi | 922 | 6 | 0.6 | 166 | 2 | AE1212 | lipoprotein signal |
| 850 | 6 | 0.6 | 152 | 2 | C75544 | hypothetical prote | 923 | 6 | 0.6 | 166 | 2 | D90020 | 30S ribosomal prot |
| 851 | 6 | 0.6 | 152 | 2 | T24064 | hypothetical prote | 924 | 6 | 0.6 | 166 | 2 | A83111 | 50S ribosomal prot |
| 852 | 6 | 0.6 | 153 | 1 | ICHU2 | interleukin-2 prec | 925 | 6 | 0.6 | 166 | 2 | A28127 | myosin light chain |
| 853 | 6 | 0.6 | 153 | 1 | ICG12 | interleukin-2 prec | 926 | 6 | 0.6 | 166 | 2 | C71874 | hypothetical prote |
| 854 | 6 | 0.6 | 153 | 2 | F75514 | conserved hypotcet | 927 | 6 | 0.6 | 167 | 1 | RDSODF | dihydrofolate redu |
| 855 | 6 | 0.6 | 153 | 2 | S47817 | hypothetical 18.1k | 928 | 6 | 0.6 | 167 | 2 | B32646 | peroxisomal membra |
| 856 | 6 | 0.6 | 153 | 2 | G91187 | hypothetical prote | 929 | 6 | 0.6 | 167 | 2 | A32646 | peroxisomal membra |
| 857 | 6 | 0.6 | 153 | 2 | H86034 | hypothetical prote | 930 | 6 | 0.6 | 167 | 2 | S65670 | phosphoprotein pho |
| 858 | 6 | 0.6 | 153 | 2 | F85642 | hypothetical prote | 931 | 6 | 0.6 | 167 | 2 | G95335 | hypothetical prote |
| 859 | 6 | 0.6 | 153 | 2 | A90782 | hypothetical prote | 932 | 6 | 0.6 | 167 | 2 | AI1129 | hypothetical prote |
| 860 | 6 | 0.6 | 153 | 2 | AF3648 | hypothetical prote | 933 | 6 | 0.6 | 168 | 2 | C96949 | probable transcrip |
| 861 | 6 | 0.6 | 153 | 2 | G86784 | hypothetical prote | 934 | 6 | 0.6 | 168 | 2 | AE0192 | probable C4-dicarb |
| 862 | 6 | 0.6 | 153 | 2 | S67294 | hypothetical prote | 935 | 6 | 0.6 | 169 | 2 | T02943 | ubiquitin-conjugat |
| 863 | 6 | 0.6 | 154 | 2 | JN0698 | interleukin 2 prec | 936 | 6 | 0.6 | 169 | 2 | T01329 | ubiquitin-conjugat |
| 864 | 6 | 0.6 | 154 | 2 | A69946 | hypothetical prote | 937 | 6 | 0.6 | 169 | 2 | C70207 | outer membrane pro |
| 865 | 6 | 0.6 | 155 | 2 | C64314 | conserved hypotcet | 938 | 6 | 0.6 | 169 | 2 | B87602 | hypothetical prote |
| 866 | 6 | 0.6 | 156 | 1 | S06078 | Ha-transporing tw | 939 | 6 | 0.6 | 170 | 2 | T03768 | disease resistance |
| 867 | 6 | 0.6 | 156 | 2 | S15266 | fimA protein - Dic | 940 | 6 | 0.6 | 170 | 2 | S59869 | TRAP-like protein |
| 868 | 6 | 0.6 | 156 | 2 | AH2247 | hypothetical prote | 941 | 6 | 0.6 | 170 | 2 | B86341 | F9H16.13 protein - |
| 869 | 6 | 0.6 | 156 | 2 | S31000 | gene 55 protein - | 942 | 6 | 0.6 | 170 | 2 | C86652 | hypothetical prote |
| 870 | 6 | 0.6 | 156 | 2 | D71132 | hypothetical prote | 943 | 6 | 0.6 | 170 | 2 | T49248 | zinc finger-like p |
| 871 | 6 | 0.6 | 157 | 1 | S47575 | EGD1 protein - yea | 944 | 6 | 0.6 | 171 | 2 | T01466 | iron-sulfur cofact |
| 872 | 6 | 0.6 | 157 | 2 | B70377 | iron-sulfur cofact | 945 | 6 | 0.6 | 171 | 2 | T27371 | peptidylprolyl iso |
| 873 | 6 | 0.6 | 157 | 2 | T09526 | stress response ge | 946 | 6 | 0.6 | 171 | 2 | B86719 | hypothetical prote |
| 874 | 6 | 0.6 | 157 | 2 | T09659 | pathogenesis-relat | 947 | 6 | 0.6 | 171 | 2 | H70527 | probable PE protei |
| 875 | 6 | 0.6 | 157 | 2 | A46566 | pilin precursor - | 948 | 6 | 0.6 | 171 | 2 | H64960 | outer membrane por |
| 876 | 6 | 0.6 | 157 | 2 | S30129 | macrogamete/sporo | 949 | 6 | 0.6 | 171 | 2 | AE2604 | conserved hypotcet |
| 877 | 6 | 0.6 | 157 | 2 | H69849 | hypothetical prote | 950 | 6 | 0.6 | 171 | 2 | E97386 | hypothetical prote |
| 878 | 6 | 0.6 | 157 | 2 | F75612 | conserved hypotcet | 951 | 6 | 0.6 | 172 | 2 | T27882 | peptidylprolyl iso |
| 879 | 6 | 0.6 | 157 | 2 | G64334 | peptidylprolyl iso | 952 | 6 | 0.6 | 172 | 2 | T06073 | peptidylprolyl iso |
| 880 | 6 | 0.6 | 157 | 2 | AG1312 | thioredoxin homolo | 953 | 6 | 0.6 | 172 | 2 | F84383 | hypothetical prote |
| 881 | 6 | 0.6 | 158 | 2 | B54762 | phospholipase A2 (| 954 | 6 | 0.6 | 172 | 2 | AF2528 | hypothetical prote |
| 882 | 6 | 0.6 | 158 | 2 | T06527 | pathogenesis-relat | 955 | 6 | 0.6 | 172 | 2 | F87649 | ExbD/ToIR family p |
| 883 | 6 | 0.6 | 158 | 2 | A75536 | hypothetical prote | 956 | 6 | 0.6 | 173 | 2 | T27373 | peptidylprolyl iso |
| 884 | 6 | 0.6 | 158 | 2 | T09470 | hypothetical prote | 957 | 6 | 0.6 | 173 | 2 | A24859 | legumin A - tick b |
| 885 | 6 | 0.6 | 158 | 2 | B85073 | hypothetical prote | 958 | 6 | 0.6 | 173 | 2 | A72862 | AcOrf-96 protein - |
| 886 | 6 | 0.6 | 159 | 2 | F82548 | hypothetical prote | 959 | 6 | 0.6 | 173 | 2 | H86631 | hypothetical prote |
| 887 | 6 | 0.6 | 159 | 2 | B83238 | hypothetical prote | 960 | 6 | 0.6 | 173 | 2 | E42526 | BeR 20.2K protein |
| 888 | 6 | 0.6 | 159 | 2 | H86255 | protein P12f1.6 [i | 961 | 6 | 0.6 | 173 | 2 | G89852 | hypothetical prote |
| 889 | 6 | 0.6 | 159 | 2 | E90327 | conserved hypotcet | 962 | 6 | 0.6 | 174 | 2 | S58492 | auxin-induced prot |
| 890 | 6 | 0.6 | 159 | 2 | T04297 | hypothetical prote | 963 | 6 | 0.6 | 174 | 2 | S12244 | auxin-induced prot |
| 891 | 6 | 0.6 | 159 | 2 | H95375 | hypothetical prote | 964 | 6 | 0.6 | 174 | 2 | G86289 | auxin-induced prot |
| 892 | 6 | 0.6 | 159 | 2 | S49966 | probable membrane | 965 | 6 | 0.6 | 174 | 2 | S16528 | hypothetical prote |
| 893 | 6 | 0.6 | 160 | 2 | R87463 | glutathione peroxi | 966 | 6 | 0.6 | 174 | 2 | D64150 | hypothetical prote |
| 894 | 6 | 0.6 | 160 | 2 | C89832 | hypothetical prote | 967 | 6 | 0.6 | 174 | 2 | F70635 | hypothetical prote |
| 895 | 6 | 0.6 | 160 | 2 | A45607 | 16k surface antige | 968 | 6 | 0.6 | 175 | 2 | T05380 | hypothetical prote |
| 896 | 6 | 0.6 | 160 | 2 | T50732 | 2-vinyl bacterioch | 969 | 6 | 0.6 | 175 | 2 | T11504 | NADH2 dehydrogenas |
| 897 | 6 | 0.6 | 161 | 1 | AFKTA | allophycocyanin be | 970 | 6 | 0.6 | 175 | 2 | C75300 | hypothetical prote |
| 898 | 6 | 0.6 | 161 | 2 | S30938 | allophycocyanin b | 971 | 6 | 0.6 | 175 | 2 | B70762 | probable fadB11' p |
| 899 | 6 | 0.6 | 161 | 2 | S03938 | neuroendocrine pro | 972 | 6 | 0.6 | 176 | 2 | I84638 | pituitary adenylat |
| 900 | 6 | 0.6 | 161 | 2 | T21405 | hypothetical prote | 973 | 6 | 0.6 | 176 | 2 | S30199 | ribosomal protein |
| 901 | 6 | 0.6 | 161 | 2 | B72551 | probable cycchrom | 974 | 6 | 0.6 | 177 | 2 | S23505 | chymase (EC 3.4.21 |
| 902 | 6 | 0.6 | 161 | 2 | AG1132 | hypothetical prote | 975 | 6 | 0.6 | 177 | 2 | S57399 | small heat shock p |
| 903 | 6 | 0.6 | 162 | 2 | T50253 | Vacuolar ATP synth | 976 | 6 | 0.6 | 177 | 2 | C97680 | hypothetical prote |
| 904 | 6 | 0.6 | 162 | 2 | C69228 | ATP synthase, subu | 977 | 6 | 0.6 | 178 | 2 | S31768 | cystathionine beta |
| 905 | 6 | 0.6 | 162 | 2 | C45731 | motB protein - pha | 978 | 6 | 0.6 | 178 | 2 | AB0870 | outer membrane pro |

979 6 0.6 178 2 A70227
980 6 0.6 179 2 G37729
981 6 0.6 179 2 S51904
982 6 0.6 179 2 S74360
983 6 0.6 179 2 B70757
984 6 0.6 179 2 H86273
985 6 0.6 180 2 H95095
986 6 0.6 180 2 D59475
987 6 0.6 181 2 I69036
988 6 0.6 181 2 T18074
989 6 0.6 181 2 A12402
990 6 0.6 181 2 E86488
991 6 0.6 181 2 S74786
992 6 0.6 181 2 E86555
993 6 0.6 182 1 RXP6
994 6 0.6 182 2 F81366
995 6 0.6 182 2 C83600
996 6 0.6 182 2 A84749
997 6 0.6 182 2 T1214
998 6 0.6 183 2 A28335
999 6 0.6 183 2 A5092
1000 6 0.6 183 2 H71894

ALIGNMENTS

RESULT 1
141078
hemolysin - Escherichia coli
C:Species: Escherichia coli
C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 21-Jul-2000
C:Accession: I41078
R:Schmidt, H.; Beutin, L.; Karch, H.
J. Infect. Immun. 63, 1055-1061, 1995
A:Title: Molecular analysis of the plasmid-encoded hemolysin of Escherichia coli O157:H7
A:Reference number: I41077; MUID: 95172699; PMID: 7868227
A:Accession: I41078
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-998 <RES>
A:Cross-references: EMBL:X79839; NID:Q860924; PIDN:CAA56234.1; PID:G4388764
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: lipoprotein
F:233-776/Domain: hemolysin A homology <HLYA>
F:550,675/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match 1.5% Score 14; DB 2; Length 998;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 379 ISGILEASKQAMFE 392
|||||
Db 392 ISGILEASKQAMFE 405

RESULT 2
T00227
hemolysin A toxin protein - Escherichia coli plasmid p0157
C:Species: Escherichia coli
C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00227; T42148
R:Makino, K.; Ishii, K.; Yasunaga, T.; Hattori, M.; Yokoyama, K.; Yatsudo, H.C.; Kubota, S.; Shinagawa, H.
DNA Res. 5, 1-9, 1998
A:Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic E. coli O157:H7
A:Reference number: 214127; MUID: 98290540; PMID: 9628576
A:Accession: T00227
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-998 <MAK>
A:Cross-references: EMBL:AB011549; NID:G4589740; PIDN:BA031774.1; PID:G3337015
A:Experimental source: strain EHEC O157:H7, substrain RIMD 0509952

R:Burland, V.; Shao, Y.; Perna, N.T.; Plunkett, G.; Sofia, H.J.; Blattner, F.R.
Nucleic Acids Res. 26, 4196-4204, 1998
A:Title: The complete DNA sequence and analysis of the large virulence plasmid of Escherichia coli O157:H7
A:Reference number: Z22068; MUID: 98391744; PMID: 9722640
A:Accession: T42148
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-998 <BUR>
A:Cross-references: EMBL:AF074613; PIDN:AACT0116.1
A:Experimental source: strain EDL933; serotype O157:H7
C:Genetics:
A:Gene: hlyA
A:Genome: plasmid p0157
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: cytotoxicity; hemolysin; lipoprotein; toxin
F:233-776/Domain: hemolysin A homology <HLYA>
F:550,675/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match 1.5% Score 14; DB 2; Length 998;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 379 ISGILEASKQAMFE 392
|||||
Db 392 ISGILEASKQAMFE 405

RESULT 3
LEECA
hemolysin A - Escherichia coli
C:Species: Escherichia coli
C>Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 18-Jun-1999
C:Accession: A24433; I41280
R:Feilmee, T.; Pellett, S.; Welch, R.A.
J. Bacteriol. 163, 94-105, 1985
A:Title: Nucleotide sequence of an Escherichia coli chromosomal hemolysin.
A:Reference number: A24433; MUID: 85234404; PMID: 3891743
A:Accession: A24433
A:Molecule type: DNA
A:Residues: 1-1023 <FEL>
A:Cross-references: GB:M10133; GB:M12863; NID:G146377; PIDN:AAA33975.1; PID:G146379
A:Experimental source: strain J96, O4 serotype
R:Stanley, P.; Packman, L.C.; Koronakis, V.; Hughes, C.
Science 266, 1992-1996, 1994
A:Title: Pasty acylation of two internal lysine residues required for the toxic activity of hemolysin A
A:Reference number: A55387; MUID: 9509325; PMID: 7801126
A:Contents: annotation; lysine palmitoylation
A:Note: lysine modification is performed by the hlyC gene product
R:Haertlein, M.; Schlessel, S.; Wagner, W.; Rdest, U.; Kref, J.; Goebel, W.
J. Cell Biol. 22, 87-97, 1983
A:Title: Transport of hemolysin by Escherichia coli.
A:Reference number: I41280
A:Accession: I41280
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1, 'T', 3, 'V', 5, 'T', 7-44 <RES>
A:Cross-references: GB:M29173; NID:G146337; PIDN:AAA23957.1; PID:G146338
C:Genetics:
A:Gene: hlyA
C:Function:
A:Description: attacks blood cell membranes and causes cell lysis
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: calcium binding; cytotoxicity; exotoxin; hemolysin; lipoprotein; tandem repeat;
F:246-791/Domain: hemolysin A homology <HLYA>
F:723-851/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVYF]-X)
F:563,689/Binding site: palmitate (Lys) (covalent) #status experimental

Query Match 1.5% Score 14; DB 1; Length 1023;
Best Local Similarity 100.0%; Pred. No. 6.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 379 ISGILEASKQAMFE 392
|||||

```

Db      405 ISGILEASKQAMPE 418

RESULT 4
hemolysin A - Escherichia coli plasmid pHy152
S10056
C:Species: Escherichia coli
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 04-Mar-2000
C:Accession: S10056
R:Hees, J.; Weis, W.; Vogel, M.; Goebel, W.
FEMS Microbiol. Lett. 34, 1-11, 1986
A:Title: Nucleotide sequence of a plasmid-encoded hemolysin determinant and its comparison with the hemolysin A gene of Escherichia coli
A:Reference number: S07209
A:Accession: S10056
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1024 <HES>
A:Cross-references: EMBL:M14107
C:Genetics:
A:Genome: plasmid pHy152
A:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: lipoprotein
F:247-792/Domain: hemolysin A homology <HLYA>
F:564,690/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match      1.5%; Score 14; DB 2; Length 1024;
Best Local Similarity 100.0%; Pred. No. 6.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      379 ISGILEASKQAMPE 392
Db      406 ISGILEASKQAMPE 419

RESULT 5
S51784
toxin III - Actinobacillus pleuropneumoniae (serotype 2)
N:Alternate names: RTX-toxin IIIA (ApXIIIA)
C:Species: Actinobacillus pleuropneumoniae
C:Date: 14-Jul-1995 #sequence_revision 15-Nov-1996 #text_change 18-Jun-1999
C:Accession: S51784
R:Chang, Y.F.; Shi, J.; Ma, D.P.; Shin, S.J.; Lein, D.H.
DNA Cell Biol. 12, 351-362, 1993
A:Title: Molecular analysis of the Actinobacillus pleuropneumoniae RTX toxin-III gene cluster
A:Reference number: S51783; MUID:93263992; PMID:8494611
A:Accession: S51784
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1049 <CHA>
A:Cross-references: EMBL:L12145; NID:g349605; PIDN:AAA21924.1; PID:g470685
C:Comment: This organism causes porcine pleuropneumonia.
C:Genetics:
A:Gene: apXIIIA
C:Function:
A:Description: lyse lung macrophages
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: calcium binding; cytolysis; exotoxin; lipoprotein; tandem repeat; thiolester
F:254-803/Domain: hemolysin A homology <HLYA>
F:735-861/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIFY]-X)
F:571,702/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match      1.5%; Score 14; DB 1; Length 1049;
Best Local Similarity 100.0%; Pred. No. 6.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      241 KVAAGFELSNOVIG 254
Db      275 KVAAGFELSNOVIG 288

RESULT 6
B49219
toxin III - Actinobacillus pleuropneumoniae (serotype 8)

```

```

N:Alternate names: RTX-toxin III (ApXIIIA)
C:Species: Actinobacillus pleuropneumoniae
C:Date: 19-Dec-1993 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C:Accession: B49219; S48043; S29958
R:Jansen, R.; Briaire, J.; Kamp, E.M.; Gielkens, A.L.; Smits, M.A.
Infect. Immun. 61, 947-954, 1993
A:Title: Cloning and characterization of the Actinobacillus pleuropneumoniae-RTX-toxin III
A:Reference number: A49219; MUID:93162836; PMID:8432615
A:Accession: B49219
A:Molecule type: DNA
A:Residues: 1-1052 <JANI>
A:Cross-references: EMBL:X68815; NID:g38956; PIDN:CAA48711.1; PID:g38958
A:Experimental source: strain 405, serotype 8
A:Note: sequence extracted from NCBI backbone (NCBIN:125168, NCBIIP:125170)
R:Jansen, R.; Briaire, J.; van Geel, A.B.M.; Kamp, E.M.; Gielkens, A.L.J.; Smits, M.A.
Infect. Immun. 62, 4411-4418, 1994
A:Title: Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (ApX) operons: characterization of the RTX-toxin gene from Actinobacillus pleuropneumoniae
A:Reference number: S48042; MUID:95012630; PMID:7927703
A:Accession: S48043
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1052 <JANI>
A:Cross-references: EMBL:X68055; NID:g558150; PIDN:CAA56358.1; PID:g558152
A:Experimental source: strain 405, serotype 8
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Comment: This organism causes porcine pleuropneumonia.
C:Genetics:
A:Gene: apXIIIA
C:Function:
A:Description: lyse lung macrophages
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: calcium binding; cytolysis; exotoxin; lipoprotein; tandem repeat; thiolester
F:254-804/Domain: hemolysin A homology <HLYA>
F:736-862/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIFY]-X)
F:571,702/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match      1.5%; Score 14; DB 1; Length 1052;
Best Local Similarity 100.0%; Pred. No. 6.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      241 KVAAGFELSNOVIG 254
Db      275 KVAAGFELSNOVIG 288

RESULT 7
B33389
toxin II - Actinobacillus pleuropneumoniae
N:Alternate names: cytolysin II; RTX-toxin II (ApXII)
C:Species: Actinobacillus pleuropneumoniae
C:Date: 09-Mar-1990 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C:Accession: B33389; S18853; B43599
R:Chang, Y.F.; Young, R.; Struck, D.K.
DNA 8, 635-647, 1989
A:Title: Cloning and characterization of a hemolysin gene from Actinobacillus (Haemophilus)
A:Reference number: A33389; MUID:90126233; PMID:2693022
A:Accession: B33389
A:Molecule type: DNA
A:Residues: 1-956 <CHA>
A:Cross-references: GB:M30602; NID:g141823; PIDN:AAA87232.1; PID:g141825
A:Experimental source: serotype 5
R:Smits, M.A.; Briaire, J.; Jansen, R.; Smith, H.E.; Kamp, E.M.; Gielkens, A.L.J.
submitted to the EMBL Data Library, July 1991
A:Description: Cytolysins of Actinobacillus pleuropneumoniae serotype 9.
A:Reference number: S18852
A:Accession: S18853
A:Molecule type: DNA
A:Residues: 1-956 <SMI>
A:Cross-references: EMBL:X61111; NID:g38939; PIDN:CAA43423.1; PID:g38941
R:Smits, M.A.; Briaire, J.; Jansen, R.; Smith, H.E.; Kamp, E.M.; Gielkens, A.L.J.
Infect. Immun. 59, 4497-4504, 1991
A:Title: Cytolysins of Actinobacillus pleuropneumoniae serotype 9.
A:Reference number: A43599; MUID:92040145; PMID:1937809

```


A;Accession: B43599
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-27;948-956 <SM2>
A;Cross-references: GB:X61111; NID:g338939
C;Comment: This organism causes porcine pleuropneumonia.
C;Genetics:
A;Gene: apXIIA; appA; clyIIA
C;Function:
A;Description: attacks blood cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;243-787/Domain: hemolysin A homology <HLXA>
F;719-801/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIFY]-X)
F;557/Binding site: palmitate (lys) (covalent) #status predicted

Query Match 1.2%; Score 11; DB 1; Length 956;
Best Local Similarity 100.0%; Pred. No. 0.062;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 369 ALLVAGVTGLI 379
|||||
Db 392 ALLVAGVTGLI 402

RESULT 8
A43834
toxin II - Actinobacillus suis
N;Alternate names: asha protein; cytolysin II; RTX-toxin II
C;Species: Actinobacillus suis
C;Date: 31-Dec-1993 #sequence_revision 08-Nov-1996 #text_change 05-Dec-1998
C;Accession: A43834
R;Burrows, L.L.; Lo, R.Y.
Infect. Immun. 60, 2166-2173, 1992
A;Title: Molecular characterization of an RTX toxin determinant from Actinobacillus suis
A;Reference number: A43834; MUID:92267623; PMID:1587585
A;Accession: A43834
A;Molecule type: DNA
A;Residues: 1-956 <BUR>
A;Experimental source: isolate 3714
A;Note: Sequence extracted from NCBI backbone (NCBIN:104212, NCBI:P:104211)
C;Comment: This organism causes acute fatal septicemia in young pigs.
C;Function:
A;Description: attacks cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;243-787/Domain: hemolysin A homology <HLXA>
F;719-801/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIFY]-X)
F;557/Binding site: palmitate (lys) (covalent) #status predicted

Query Match 1.2%; Score 11; DB 1; Length 956;
Best Local Similarity 100.0%; Pred. No. 0.062;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 369 ALLVAGVTGLI 379
|||||
Db 392 ALLVAGVTGLI 402

RESULT 9
I39643
RTX-toxin I - Actinobacillus pleuropneumoniae
N;Alternate names: hemolysin ApXI
C;Species: Actinobacillus pleuropneumoniae
C;Date: 19-Jul-1996 #sequence_revision 08-Nov-1996 #text_change 18-Jun-1999
C;Accession: I39643; S18769; I39645; S60732; S35781
R;Jansen, R.; Briaire, J.; Kamp, E.M.; Gielkens, A.L.; Smits, M.A.
Infect. Immun. 61, 3688-3695, 1993
A;Title: Structural analysis of the Actinobacillus pleuropneumoniae RTX-toxin I (ApXI)
A;Reference number: I39641; MUID:93366425; PMID:8359891
A;Accession: I39643
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
A;Molecule type: DNA

A;Residues: 1-1022 <RES>
A;Cross-references: EMBL:X73117; NID:g312897; PIDN:CAA51548.1; PID:g312899
R;Frey, J.; Meier, R.; Gysi, D.; Nicolet, J.
Infect. Immun. 59, 3026-3032, 1991
A;Title: Nucleotide sequence of the hemolysin I gene from Actinobacillus pleuropneumoniae
A;Reference number: S18769; MUID:91348845; PMID:1879928
A;Accession: S18769
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-209, 'AMPYLTLA', 218-373, 'R', 375-561, 'Q', 563-686, 'TC', 688-1022 <PRE>
A;Cross-references: EMBL:X52899; NID:g38949; PIDN:CAA37081.1; PID:g38950
R;Frey, J.; Haldmann, A.; Nicolet, J.; Boffini, A.; Prentki, P.
Gene 142, 97-102, 1994
A;Title: Sequence analysis and transcription of the apxI operon (hemolysin I) from Actinobacillus pleuropneumoniae
A;Reference number: I39644; MUID:94237497; PMID:8181764
A;Accession: I39645
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
A;Molecule type: DNA
A;Residues: 1-209, 'AMPYLTLA', 218-373, 'R', 375-561, 'Q', 563-686, 'TC', 688-1022 <RE2>
A;Cross-references: EMBL:X68595; NID:g505568; PIDN:CAA48586.1; PID:g505570
R;Tascón, R.I.; Vazquez-Boland, J.A.; Gutierrez-Martin, C.B.; Rodriguez-Barbosa, I.; Roda-Navarro, J.
Mol. Microbiol. 14, 207-216, 1994
A;Title: The RTX haemolysins Apxi and ApxiI are major virulence factors of the swine pathogen Actinobacillus pleuropneumoniae
A;Reference number: S60731; MUID:95131743; PMID:7830567
A;Accession: S60732
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 604-685 <TAS>
C;Comment: This organism causes porcine pleuropneumonia.
C;Genetics:
A;Gene: apXIA
C;Function:
A;Description: attacks blood cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;243-789/Domain: hemolysin A homology <HLXA>
F;721-847/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIFY]-X)
F;560,686/Binding site: palmitate (lys) (covalent) #status predicted

Query Match 1.2%; Score 11; DB 1; Length 1022;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 266 AQRVAGLSTT 276
|||||
Db 289 AQRVAGLSTT 299

RESULT 10
B30169
leukotoxin A - Pasteurella haemolytica (serotype 1)
N;Alternate names: lktA protein
C;Species: Pasteurella haemolytica
C;Date: 12-Oct-1989 #sequence_revision 15-Nov-1996 #text_change 18-Jun-1999
C;Accession: B30169; C32051; S29516
R;Highlander, S.K.; Chidambaram, M.; Engler, M.J.; Weinstock, G.M.
DNA 8, 15-28, 1989
A;Title: DNA sequence of the Pasteurella haemolytica leukotoxin gene cluster.
A;Reference number: A30169; MUID:89210283; PMID:2707120
A;Accession: B30169
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-953 <HIG>
R;Strathdee, C.A.; Lo, R.Y.C.
J. Bacteriol. 171, 916-928, 1989
A;Title: Cloning, nucleotide sequence, and characterization of genes encoding the secreted leukotoxin of Pasteurella haemolytica
A;Reference number: A32051; MUID:89123172; PMID:2914876
A;Accession: C32051
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 947-953 <STR>
R;Lo, R.Y.C.; Strathdee, C.A.; Shewen, P.E.
Infect. Immun. 55, 1987-1996, 1987

A;Title: Nucleotide sequence of the leukotoxin genes of *Pasteurella haemolytica* A1.
A;Reference number: S29515; MUID:87306837; PMID:3040588
A;Accession: S29516
A;Molecule type: DNA
A;Residues: 1-741,'D',743-953 <LOR>
A;Cross-references: EMBL:M20730; NID:g150492; PIDN:AAA25529.1; PID:g150494
C;Comment: This organism causes bovine pneumonic pasteurellosis (shipping fever).
C;Genetics:
A;Gene: lktA

A;Function: lysates leukocytes
A;Description: hemolysin A; hemolysin A homology
C;Superfamily: hemolysin A; hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;238-784/Domain: hemolysin A homology <HLA>
F;716-807/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVYF]-X)
F;554/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match 1.1%; Score 10; DB 1; Length 953;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAORVAAGLS 274
|||||
Db 283 LAORVAAGLS 292

RESULT 11
A35254
leukotoxin A - *Pasteurella haemolytica* (serotype T10)
N;Alternate names: lktA protein
C;Species: *Pasteurella haemolytica*
C;Date: 10-Aug-1990 #sequence revision 15-Nov-1996 #text_change 18-Jun-1999
C;Accession: S37145; A35254; S34237; S34235
R;Lainson, A.F.; Aitchison, K.; Donachie, W.
submitted to the EMBL Data Library, September 1993
A;Description: DNA sequence of the leukotoxin A gene from *P. haemolytica* T10 serotype.
A;Reference number: S37145

A;Accession: S37145
A;Molecule type: DNA
A;Residues: 1-955 <LA1>
A;Cross-references: EMBL:Z26247; NID:g400424; PIDN:CAA81206.1; PID:g400425
R;Highlander, S.K.; Engler, M.J.; Weinstein, G.M.
J. Bacteriol. 172, 2343-2350, 1990
A;Title: Secretion and expression of the *Pasteurella haemolytica* leukotoxin.
A;Reference number: A35254; MUID:90236888; PMID:2185213

A;Accession: A35254
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 950-955 <HIG>
A;Cross-references: GB:M24197; GB:M34943; GB:M34944
R;Lainson, A.F.; Aitchison, K.D.; Donachie, W.
submitted to the EMBL Data Library, June 1993
A;Description: DNA sequence of the carboxy terminal end of leukotoxin A from the T3 sero

A;Reference number: S34237
A;Accession: S34235
A;Molecule type: DNA
A;Residues: 745-955 <LA2>
A;Cross-references: EMBL:Z22884; NID:g311828; PIDN:CAA80498.1; PID:g311829
A;Experimental source: serotype T3

A;Accession: S34235
A;Molecule type: DNA
A;Residues: 723-955 <LA3>
A;Cross-references: EMBL:Z22887; NID:g311824; PIDN:CAA80501.1; PID:g311825
A;Experimental source: serotype T10
C;Function:

A;Description: attacks cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;240-786/Domain: hemolysin A homology <HLA>
F;718-809/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVYF]-X)
F;718-726/Region: repeat
F;727-735/Region: repeat
F;736-744/Region: repeat

F;745-753/Region: repeat
F;754-762/Region: repeat
F;763-771/Region: repeat
F;772-780/Region: repeat
F;781-789/Region: repeat
F;792-800/Region: repeat
F;801-809/Region: repeat
F;556/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match 1.1%; Score 10; DB 1; Length 955;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAORVAAGLS 274
|||||
Db 285 LAORVAAGLS 294

RESULT 12
S77624
mannuronan C-5-epimerase (EC 5.1.3.-) - *Azotobacter vinelandii*
C;Species: *Azotobacter vinelandii*
C;Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 21-Jul-2000
C;Accession: S77624
R;Ertesvaag, H.; Hoidal, H.K.; Hals, I.K.; Rian, A.; Doseeth, B.; Valla, S.
Mol. Microbiol. 16, 719-731, 1995
A;Title: A family of modular type mannuronan C-5-epimerase genes controls alginate struct
A;Reference number: I39738; MUID:96065700; PMID:7476166
A;Accession: S77624
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-1403 <ERT>
A;Cross-references: EMBL:L39096; NID:g790690; PIDN:AAA87311.1; PID:g790692
A;Experimental source: strain E
C;Genetics:
A;Gene: algE1
C;Function:

A;Description: catalyzes the Ca(2+)-dependent epimerization of D-mannuronic acid residues
A;Pathway: alginate biosynthesis
C;Keywords: calcium binding; isomerase
Query Match 1.1%; Score 10; DB 2; Length 1403;
Best Local Similarity 100.0%; Pred. No. 0.88;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 714 GDLLDGAG 723
|||||
Db 1257 GDLLDGAG 1266

RESULT 13
AH2515
hypothetical protein alr7304 [imported] - *Nostoc* sp. (strain PCC 7120) plasmid pCC7120alt
C;Species: *Nostoc* sp. PCC 7120
A;Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AH2515
R;kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing *Cyanobacterium* Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2515
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-4936 <KUR>
A;Cross-references: GB:BA000020; PIDN:BA078388.1; PID:g17135842; GSPDB:GN00180
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr7304
A;Genome: plasmid

Query Match 1.1%; Score 10; DB 2; Length 4936;

Best Local Similarity 100.0%; Pred. No. 2.7; Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 752 DVLNGGAGND 761
|||||

Db 4296 DVLNGGAGND 4305

RESULT 14

AG2664

A:ATP synthase B chain atpF [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C:Accession: AG2664

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, G.; Romero, P.; Zhang, S.

A:Description: DNA sequence of the carboxy terminal end of leukotoxin A from the T4 serotype 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:21608550; PMID:11743193

A:Accession: AG2664

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-161 <KUR>

A:Cross-references: GB:AB008688; PIDN:AA41733.1; PID:gl7739083; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: atpF

A:Map position: circular chromosome

Query Match 0.9%; Score 8; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 13; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 325 HLLAEYQR 332
|||||

Db 56 HLLAEYQR 63

RESULT 15

F97446

A:hypothetical protein AGR_C_1301 [imported] - Agrobacterium tumefaciens (strain C58, Cere)

C:Species: Agrobacterium tumefaciens

C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002

C:Accession: F97446

R:Goodner, B.; Hinkle, G.; Gattling, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens

A:Reference number: A97359; PMID:21608551; PMID:11743194

A:Accession: F97446

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-161 <KUR>

A:Cross-references: GB:AB007869; PIDN:AAK86527.1; PID:gl5155685; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C_1301

A:Map position: circular chromosome

Query Match 0.9%; Score 8; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 13; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 325 HLLAEYQR 332
|||||

Db 56 HLLAEYQR 63

RESULT 16

S34238

A:leukotoxin A - Pasteurella haemolytica (fragment)

N:Alternate names: lktA protein

C:Species: Pasteurella haemolytica

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Mar-2001

C:Accession: S34238; S34236

R:Lainson, A.F.; Aitchison, K.D.; Donachie, W.

A:Submitted to the EMBL Data Library, June 1993

A:Description: DNA sequence of the carboxy terminal end of leukotoxin A from the T4 serotype of Pasteurella haemolytica.

A:Reference number: S34236

A:Accession: S34238

A:Molecule type: DNA

A:Residues: 1-208 <LA1>

A:Cross-references: EMBL:Z22885; NID:g311830; PIDN:CAA80499.1; PID:g311831

A:Experimental source: serotype T4

A:Accession: S34236

A:Molecule type: DNA

A:Residues: 9-208 <LA2>

A:Cross-references: EMBL:Z22886; NID:g311826; PIDN:CAA80500.1; PID:g311827

A:Experimental source: serotype T15

C:Genetics:

A:Gene: lktA

A:Superfamily: hemolysin A; hemolysin A homology

C:Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat

Query Match 0.9%; Score 8; DB 2; Length 208;
Best Local Similarity 100.0%; Pred. No. 16; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 725 DRLFGGKG 732
|||||

Db 3 DRLFGGKG 10

RESULT 17

AD0468

A:conserved hypothetical protein YPO3844 [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001

C:Accession: AD0468

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; PMID:21470413; PMID:11586360

A:Accession: AD0468

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-234 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC93312.1; PID:gl5981759; GSPDB:GN00175

C:Genetics:

A:Gene: YPO3844

Query Match 0.9%; Score 8; DB 2; Length 234;
Best Local Similarity 100.0%; Pred. No. 18; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 134 ALAGIELD 141
|||||

Db 8 ALAGIELD 15

RESULT 18

E70714

A:hypothetical protein Rv1514c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: E70714

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70714
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-262 <COL>
A;Cross-references: GB:279701; GB:AL123456; NID:g3261635; PIDN:CAB02028.1; PID:e264143;
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: RV1514c

Query Match 0.9%; Score 8; DB 2; Length 262;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 746 DGGSGDDV 753
|||||
Db 42 DGGSGDDV 49

RESULT 19
AC2987
conserved hypothetical protein Atu3500 [imported] - Agrobacterium tumefaciens (strain C58)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AC2987
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Kreepan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AC2987
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-270 <KUR>
A;Cross-references: PIDN:AAL44313.1; PID:g17741903; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu3500
A;Map position: linear chromosome

Query Match 0.9%; Score 8; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 LIKKGDA 150
|||||
Db 142 LIKKGDA 149

RESULT 20
E98296
hypothetical protein AGR_L_2657 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C;Accession: E98296
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: E98296
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-270 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK89895.1; PID:g15159842; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_2657
A;Map position: linear chromosome

Query Match 0.9%; Score 8; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 LIKKGDA 150
|||||
Db 142 LIKKGDA 149

RESULT 21
JE0183
chitinase (EC 3.2.1.14) 1 - cone shell (Conus tulipa)
N;Alternate names: TBC-1
C;Species: Conus tulipa (tulip cone)
C;Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 07-May-1999
C;Accession: JE0183
R;Yamagami, T.; Ishiguro, M.
BioSci. Biotechnol. Biochem. 62, 1253-1257, 1998
A;Title: Complete amino acid sequences of chitinase-1 and -2 from bulbs of genus Tulipa.
A;Reference number: JE0183; MUID:98357241; PMID:9692212
A;Accession: JE0183
A;Molecule type: protein
A;Residues: 1-275 <YAM>
C;Comment: This enzyme catalyzes the hydrolysis of beta-1,4-linked homopolymers or oligo
C;Superfamily: alcohol sulfotransferase
C;Keywords: glycosidase; hydrolase

Query Match 0.9%; Score 8; DB 2; Length 275;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 692 IGSQFNDV 699
|||||
Db 7 IGSQFNDV 14

RESULT 22
JC7335
chitinase (EC 3.2.1.14) 1 - cone shell (Conus tulipa)
C;Species: Conus tulipa (tulip cone)
C;Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 15-Sep-2000
C;Accession: JC7335
R;Yamagami, T.; Tsutsuami, K.; Ishiguro, M.
Biosci. Biotechnol. Biochem. 64, 1394-1401, 2000
A;Title: Cloning, sequencing, and expression of the tulip bulb chitinase-1 cDNA.
A;Reference number: JC7335
A;Accession: JC7335
A;Molecule type: mRNA
A;Residues: 1-314 <YAM>
A;Cross-references: DDBJ:AB035668
C;Comment: This enzyme, a member of class III plant chitinases, which catalyzes the hydr
or protection against fungal pathogens.
C;Genetics:
A;Gene: tbc-1
C;Superfamily: alcohol sulfotransferase
C;Keywords: glycosidase; hydrolase

Query Match 0.9%; Score 8; DB 2; Length 314;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 692 IGSQFNDV 699
|||||
Db 33 IGSQFNDV 40

RESULT 23
AD3003
hypothetical protein accA [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AD3003
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.
 A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A:Reference number: AB2577; MUID:21609550; PMID:11743193
 A:Accession: AD3003
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-317 <KUR>
 A:Cross-references: GB:AB008689; PIDN:AAAL44442.1; PID:g17742043; GSPDB:GN00187
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: accA
 A:Map position: linear chromosome
 C:Superfamily: acetyl-CoA carboxylase, carboxyltransferase alpha chain

| | | | | |
|-----------------------|--------|-----------------|---------------|-------------------|
| Query Match | 0.9% | Score 8; | DB 2; | Length 317; |
| Best Local Similarity | 100.0% | Pred. No. 23; | | |
| Matches | 8; | Conservative 0; | Mismatches 0; | Indels 0; Gaps 0; |

Qy 869 ELKKLADE 876
 |||||
 Db 20 ELKKLADE 27

RESULT 24
 E98280
 mp8A protein [imported] - *Agrobacterium tumefaciens* (strain C58, Cereon)
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
 C:Accession: E98280
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens*
 A:Reference number: A97359; MUID:21609551; PMID:11743194
 A:Accession: E98280
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-317 <KUR>
 A:Cross-references: GB:AB007870; PIDN:AAK89767.1; PID:g15159691; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L_2394
 A:Map position: linear chromosome
 C:Superfamily: acetyl-CoA carboxylase, carboxyltransferase alpha chain

| | | | | |
|-----------------------|--------|-----------------|---------------|-------------------|
| Query Match | 0.9% | Score 8; | DB 2; | Length 317; |
| Best Local Similarity | 100.0% | Pred. No. 23; | | |
| Matches | 8; | Conservative 0; | Mismatches 0; | Indels 0; Gaps 0; |

Qy 869 ELKKLADE 876
 |||||
 Db 20 ELKKLADE 27

RESULT 25
 AB1054
 probable membrane protein ytfF [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhimurium
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
 A:Note: This species has also been called *Salmonella typhi*
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AB1054
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhimurium
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AB1054
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-321 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD06879.1; PID:g16505527; GSPDB:GN00176
 C:Genetics:
 A:Gene: ytfF

| | | | | |
|-----------------------|--------|-----------------|---------------|-------------------|
| Query Match | 0.9% | Score 8; | DB 2; | Length 321; |
| Best Local Similarity | 100.0% | Pred. No. 24; | | |
| Matches | 8; | Conservative 0; | Mismatches 0; | Indels 0; Gaps 0; |

Qy 368 IALLVAGV 375
 |||||
 Db 295 IALLVAGV 302

RESULT 26
 C87464
 lipoic acid synthase [imported] - *Caulobacter crescentus*
 C:Species: *Caulobacter crescentus*
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C:Accession: C87464
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.B.; Eisen, J.; Heidelberg, J.; B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.; Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: C87464
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-325 <SPO>
 A:Cross-references: GB:AE005673; NID:gl3423153; PIDN:AAK23711.1; GSPDB:GN00148
 C:Genetics:
 A:Gene: CCI735
 C:Superfamily: lipoic acid synthase

| | | | | |
|-----------------------|--------|-----------------|---------------|-------------------|
| Query Match | 0.9% | Score 8; | DB 2; | Length 325; |
| Best Local Similarity | 100.0% | Pred. No. 24; | | |
| Matches | 8; | Conservative 0; | Mismatches 0; | Indels 0; Gaps 0; |

Qy 715 DDLDDGGA 722
 |||||
 Db 135 DDLDDGGA 142

RESULT 27
 E95914
 probable secreted calcium-binding protein [imported] - *Sinorhizobium meliloti* (strain 1021)
 C:Species: *Sinorhizobium meliloti*
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: E95914
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan, P.; Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endosymbiont of the legume *Sinorhizobium meliloti*.
 A:Reference number: A95942; MUID:21396508; PMID:11481431
 A:Accession: E95914
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-353 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CAC48981.1; PID:g15140466; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid pSymB
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, H.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.; A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 C:Contents: annotation
 C:Genetics:
 A:Gene: SMB20838
 A:Genome: plasmid

Query Match 0.9%; Score 8; DB 2; Length 353;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 LDGSGDD 752
 |||||
 DB 221 LDGSGDD 228

RESULT 28
 AF1434
 AA3-600 quinol oxidase chain II [imported] - Listeria innocua (strain Clp11262)
 C;Species: Listeria innocua
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 03-Jun-2002
 C;Accession: AF1434
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A;Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
 ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
 A;Title: Comparative genomics of Listeria species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: AF1434
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-368 <GLA>
 A;Cross-references: GB:AL592022; PIDN:CAC95246.1; PID:g16412434; GSPDB:GN00178
 A;Experimental source: strain Clp11262
 C;Genetics:
 A;Gene: qoxA
 C;Superfamily: bo-type ubiquinol oxidase chain II precursor; cytochrome-c oxidase chain
 C;Keywords: copper; electron transfer; membrane-associated complex; respiratory chain

Query Match 0.9%; Score 8; DB 2; Length 368;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 374 GVTGLISG 381
 |||||
 DB 15 GVTGLISG 22

RESULT 29
 AF1806
 AA3-600 quinol oxidase chain II [imported] - Listeria monocytogenes (strain EGD-e)
 C;Species: Listeria monocytogenes
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 03-Jun-2002
 C;Accession: AF1806
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A;Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
 ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
 A;Title: Comparative genomics of Listeria species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: AF1806
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-368 <GLA>
 A;Cross-references: GB:NC_003210; PIDN:CAC98228.1; PID:g16409372; GSPDB:GN00177
 A;Experimental source: strain EGD-e
 C;Genetics:
 A;Gene: qoxA
 C;Superfamily: bo-type ubiquinol oxidase chain II precursor; cytochrome-c oxidase chain
 C;Keywords: copper; electron transfer; membrane-associated complex; respiratory chain

Query Match 0.9%; Score 8; DB 2; Length 368;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 374 GVTGLISG 381
 |||||
 DB 15 GVTGLISG 22

RESULT 29
 AF1806
 AA3-600 quinol oxidase chain II [imported] - Listeria monocytogenes (strain EGD-e)
 C;Species: Listeria monocytogenes
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 03-Jun-2002
 C;Accession: AF1806
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A;Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
 ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
 A;Title: Comparative genomics of Listeria species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: AF1806
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-368 <GLA>
 A;Cross-references: GB:NC_003210; PIDN:CAC98228.1; PID:g16409372; GSPDB:GN00177
 A;Experimental source: strain EGD-e
 C;Genetics:
 A;Gene: qoxA
 C;Superfamily: bo-type ubiquinol oxidase chain II precursor; cytochrome-c oxidase chain
 C;Keywords: copper; electron transfer; membrane-associated complex; respiratory chain

QY 374 GVTGLISG 381
 |||||
 DB 15 GVTGLISG 22

RESULT 30

AB3160
 saothpine deaminating protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont
 C;Species: Agrobacterium tumefaciens
 C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C;Accession: AB3160
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan,
 .; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, F
 ster, E.W.
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A;Reference number: AB2577; MUID:21608550; PMID:11743193
 A;Accession: AB3160
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-380 <KUR>
 A;Cross-references: GB:AE008687; PIDN:AA45696.1; PID:g17743424; GSPDB:GN00188
 A;Experimental source: strain C58 (Dupont)
 C;Genetics:
 A;Gene: agaE
 A;Genome: plasmid

Query Match 0.9%; Score 8; DB 2; Length 380;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 365 GTPIALLV 372
 |||||
 DB 230 GTPIALLV 237

RESULT 31

T19181
 hypothetical protein C10C5.5 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T19181
 R;Matthews, P.
 submitted to the EMBL Data Library, December 1995
 A;Reference number: Z19085
 A;Accession: T19181
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-394 <WIL>
 A;Cross-references: EMBL:Z68214; PIDN:CAA92446.1; GSPDB:GN00022; CESP:C10C5.5
 A;Experimental source: clone C10C5
 C;Genetics:
 A;Gene: CESP:C10C5.5
 A;Map position: 4
 A;Introns: 27/1; 48/3; 215/3; 268/1; 303/3; 330/2; 347/3

Query Match 0.9%; Score 8; DB 2; Length 394;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 ADELGIAR 55
 |||||
 DB 37 ADELGIAR 44

RESULT 32

T19180
 hypothetical protein C10C5.3 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T19180

R;Matthews, P.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z19085
A;Accession: T19180
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-399 <WIL>
A;Cross-references: EMBL:Z68214; PIDN:CAA92445.1; GSPDB:GNO0022; CESP:C10C5.3
A;Experimental source: clone C10C5
C;Genetics:
A;Gene: CESP:C10C5.3
A;Map position: 4
A;Introns: 27/1; 48/3; 215/3; 268/1; 303/3; 330/2; 350/3; 383/1

Query Match 0.9%; Score 8; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 48 ADELGIAR 55
|||||
Db 37 ADELGIAR 44

RESULT 33
E64242
GTP-binding protein obg - Mycoplasma genitalium
C;Species: Mycoplasma genitalium
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 19-Jan-2001
C;Accession: E64242
R;Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
M.; Fuhmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.
C.A.; Venter, J.C.
Science 270, 397-403, 1995
A;Title: The minimal gene complement of Mycoplasma genitalium.
A;Reference number: A64200; MUID:96026346; PMID:7569993
A;Accession: E64242
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-433 <TIGR>
A;Cross-references: GB:U99723; GB:I43967; NID:gl046092; PID:gl046095; TIGR:MG384
A;Experimental source: strain G-37
C;Genetics:
A;Genetic code: GTC3
C;Superfamily: GTP-binding protein obg; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;160-285/Domain: translation elongation factor Tu homology <ETU>
F;166-173/Region: nucleotide-binding motif A (P-loop)
F;189-194/Region: GTP-binding #status predicted
F;212-215/Region: GTP-binding #status predicted
F;282-285/Region: GTP-binding #status predicted
F;309-313/Region: GTP-binding #status predicted

Query Match 0.9%; Score 8; DB 1; Length 433;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 90 KLEFLQK 97
|||||
Db 295 KLEFLQK 302

RESULT 34
AE0826
probable cadaverine/lysine antiporter [imported] - Salmonella enterica subsp. enterica s
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: This species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AE0826
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
Th, T.; Connercon, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AE0826
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-443 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD02761.1; PID:gl6503771; GSPDB:GN00176
C;Genetics:
A;Gene: STY2805
C;Superfamily: L-lysine transport protein

Query Match 0.9%; Score 8; DB 2; Length 443;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 LGTALAGI 138
|||||
Db 226 LGTALAGI 233

RESULT 35
C84306
hypothetical protein trp6 [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: C84306
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: C84306
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-447 <STO>
A;Cross-references: GB:AE004437; NID:gl0581015; PIDN:AAG19815.1; GSPDB:GN00138
C;Genetics:
A;Gene: trp6
C;Superfamily: gamma-aminobutyric acid transporter

Query Match 0.9%; Score 8; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 358 AAVGSAVG 365
|||||
Db 16 AAVGSAVG 23

RESULT 36
H95369
EgIC ENDO-1,3-1,4-BETA-GLYCANAse (EC 3.2.1.-) [imported] - Sinorhizobium meliloti (strai
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: H95369
R;Barnett, W.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Welle, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: H95369
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-465 <KUR>
A;Cross-references: GB:AE006469; PIDN:AAK65522.1; PID:gl4523397; GSPDB:GN00165
A;Experimental source: strain 1021, megaplasmid pSYMA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Publer, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chaim, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: egIC
A:Genome: plasmid
C:Keywords: glycosidase; hydrolase

Query Match 0.9%; Score 8; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 755 NGGAGNDV 762
DB 123 NGGAGNDV 130
|||||

RESULT 37
B84372
DNA damage-inducible protein [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: B84372
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: B84372
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-469 <STO>
A:Cross-references: GB:AE004437; NID:g10581629; PIDN:AAG20342.1; GSPDB:GN00138
C:Genetics:
A:Gene: dinF

Query Match 0.9%; Score 8; DB 2; Length 469;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 357 AAAGSAV 364
DB 296 AAAGSAV 303
|||||

RESULT 38
S60902
CDP-ribitol pyrophosphorylase - Haemophilus influenzae
C:Species: Haemophilus influenzae
C:Date: 27-Apr-1996 #sequence_revision 17-May-1996 #text_change 08-Oct-1999
C:Accession: S60902; S49238; S44071
R:van Eldere, J.; Brophy, L.; Loynds, B.; Celis, P.; Hancock, I.; Carman, S.; Kroll, J.S
Mol. Microbiol. 15, 107-118, 1995
A:Title: Region II of the Haemophilus influenzae type b capsulation locus is involved in
A:Reference number: S60902; MUID:95272382; PMID:7752885
A:Accession: S60902
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-474 <VAN>
A:Cross-references: EMBL:X78559; NID:9471233; PIDN:CA55303.1; PID:g471234
A:Experimental source: serotype b; strain RMI35
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, April 1994
R:Cellis, P.; Moxon, E.; Eysen, H.; van Eldere, J.
submitted to the EMBL Data Library, September 1994
A:Description: Genetic analysis of the region II of the Haemophilus influenzae serotype
A:Reference number: S49238
A:Accession: S49238
A:Molecule type: DNA
A:Residues: 1, 'L', 3-14, 'I', 16-39, 'IF', 42-70, 'AG', 73-101, 'R', 103-212, 'F', 214-272, 'V', 274-
A:Cross-references: EMBL:Z37516; NID:g547510; PIDN:CA85750.1; PID:g547511

A:Experimental source: serotype a; strain RM107; ATCC 9006

Query Match 0.9%; Score 8; DB 2; Length 474;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 560 SKLDFSKV 567
DB 69 SKLDFSKV 76
|||||

RESULT 39
D84306

sodium- and chloride-dependent transporter [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: D84306

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: D84306
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-477 <STO>
A:Cross-references: GB:AE004437; NID:g10581016; PIDN:AAG19816.1; GSPDB:GN00138
C:Genetics:
A:Gene: nac
C:Superfamily: gamma-aminobutyric acid transporter

Query Match 0.9%; Score 8; DB 2; Length 477;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 AAVGSAGV 365
DB 15 AAVGSAGV 22
|||||

RESULT 40
G84079

sodium-dependent transporter BH3439 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G84079

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: G84079
A>Status: preliminary
A:Molecule type: DNA

A:Residues: 1-505 <STO>
A:Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BA807158.1; GSPDB:GN001

A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3439

C:Superfamily: gamma-aminobutyric acid transporter

Query Match 0.9%; Score 8; DB 2; Length 505;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 AAVGSAGV 365
DB 14 AAVGSAGV 21
|||||

Search completed: February 17, 2004, 10:21:54
Job time : 57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:13:55 ; Search time 17 Seconds

(without alignments)
2564.339 Million cell updates/sec

Title: US-10-069-799-5

Perfect score: 927

Sequence: 1 MSNINVIKSIQAGLNSTKS.....SSNALQPIPTQGTGILAPSV 927

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 14 | 1.5 | 1023 | 1 HLY1_ECOLI | P09983 escherichia |
| 2 | 14 | 1.5 | 1024 | 1 HLYA_ECOLI | P08715 escherichia |
| 3 | 14 | 1.5 | 1049 | 1 RT31_ACTPL | P55130 actinobacil |
| 4 | 14 | 1.5 | 1052 | 1 RT32_ACTPL | P55131 actinobacil |
| 5 | 11 | 1.2 | 956 | 1 HLYA_ACTSU | Q00951 actinobacil |
| 6 | 11 | 1.2 | 956 | 1 RT2A_ACTPL | P15377 actinobacil |
| 7 | 11 | 1.2 | 1023 | 1 RT12_ACTPL | P55128 actinobacil |
| 8 | 11 | 1.2 | 1023 | 1 RT12_ACTPL | P55129 actinobacil |
| 9 | 10 | 1.1 | 947 | 1 LKTA_PASSP | P55123 pasteurella |
| 10 | 10 | 1.1 | 953 | 1 LKAI_PASHA | P16535 pasteurella |
| 11 | 10 | 1.1 | 953 | 1 LKAB_PASHA | P55116 pasteurella |
| 12 | 10 | 1.1 | 953 | 1 LKAB_PASHA | P55118 pasteurella |
| 13 | 10 | 1.1 | 955 | 1 LKAA_PASHA | P55117 pasteurella |
| 14 | 8 | 0.9 | 262 | 1 YF14_MYCTU | P71793 mycobacteri |
| 15 | 8 | 0.9 | 325 | 1 LIPA_CAUCR | Q9a718 caulobacter |
| 16 | 8 | 0.9 | 433 | 1 Y384_MYCGE | P47624 mycoplasma |
| 17 | 8 | 0.9 | 465 | 1 EGLC_RHIME | Q923Q2 rhizobium m |
| 18 | 8 | 0.9 | 773 | 1 MAK5_YEAST | P38112 saccharomyc |
| 19 | 8 | 0.9 | 864 | 1 CHEA_BORBU | Q44737 borrelia bu |
| 20 | 8 | 0.9 | 1050 | 1 LKTA_ACTAC | P16462 actinobacil |
| 21 | 8 | 0.9 | 1050 | 1 CYAA_BORBR | Q57506 bordetella |
| 22 | 8 | 0.9 | 1705 | 1 RL24_PROVU | P20032 proteus vul |
| 23 | 7 | 0.8 | 63 | 1 CECA_BOMMO | Q27239 bombyx mori |
| 24 | 7 | 0.8 | 80 | 1 SCX1_TITSE | P01496 tityx serr |
| 25 | 7 | 0.8 | 80 | 1 SCX4_TITSE | P45659 tityx serr |
| 26 | 7 | 0.8 | 102 | 1 TRAC_RHISN | P55419 rhizobium s |
| 27 | 7 | 0.8 | 103 | 1 RL24_ECOLI | P02425 escherichia |
| 28 | 7 | 0.8 | 103 | 1 RL24_HAEIN | P44362 haemophilus |
| 29 | 7 | 0.8 | 103 | 1 RL24_PASHU | Q9c141 pasteurella |
| 30 | 7 | 0.8 | 104 | 1 RL24_BUCAR | P46177 buchnera ap |
| 31 | 7 | 0.8 | 111 | 1 NU4M_CAICR | Q34076 calman croc |
| 32 | 7 | 0.8 | 122 | 1 RL7_NEILA | Q9etv2 neisseria l |
| 33 | 7 | 0.8 | 122 | 1 RL7_NEIMA | P80716 neisseria m |

| | | | | | |
|-----|---|-----|-----|---------------|---------------------|
| 34 | 7 | 0.8 | 122 | 1 RL7_YERPE | Q8zap4 versinia pe |
| 35 | 7 | 0.8 | 123 | 1 RL7_NEIPE | Q9f5m1 neisseria p |
| 36 | 7 | 0.8 | 124 | 1 RL7_RALSO | Q8xuz7 ralstonia s |
| 37 | 7 | 0.8 | 125 | 1 H2AV_STRPU | P08991 strongyloce |
| 38 | 7 | 0.8 | 127 | 1 H2AV_CHICK | P02272 gallus gall |
| 39 | 7 | 0.8 | 127 | 1 H2AZ_HUMAN | P17317 homo sapien |
| 40 | 7 | 0.8 | 127 | 1 RL20_BIFLO | Q89411 bifidobacte |
| 41 | 7 | 0.8 | 131 | 1 BACH_HALAR | Q53461 haloarcula |
| 42 | 7 | 0.8 | 140 | 1 H2AV_DROME | P08985 drosophila |
| 43 | 7 | 0.8 | 141 | 1 MR2A_MYCPN | P75467 mycoplasma |
| 44 | 7 | 0.8 | 168 | 1 VTU2_DROME | P13238 drosophila |
| 45 | 7 | 0.8 | 182 | 1 ATPF_RHURU | P15013 rhodospiril |
| 46 | 7 | 0.8 | 186 | 1 NURM_NEUCR | P42116 neurospora |
| 47 | 7 | 0.8 | 187 | 1 MT28_YEAST | P40573 saccharomyc |
| 48 | 7 | 0.8 | 199 | 1 EFTS_THEMEA | Q9xtu1 thermotoga |
| 49 | 7 | 0.8 | 206 | 1 YIGZ_HAEIN | P44842 haemophilus |
| 50 | 7 | 0.8 | 208 | 1 SXIG_VOLCA | P08471 volvox cart |
| 51 | 7 | 0.8 | 210 | 1 AAC2_MYCSM | P94968 mycobacteri |
| 52 | 7 | 0.8 | 215 | 1 CYB6_CYACA | Q9t1z7 cyanidium c |
| 53 | 7 | 0.8 | 218 | 1 P25A_BOVIN | Q27957 bos taurus |
| 54 | 7 | 0.8 | 219 | 1 P25A_HUMAN | Q94811 homo sapien |
| 55 | 7 | 0.8 | 220 | 1 MT04_BACHD | Q9k7u8 bacillus ha |
| 56 | 7 | 0.8 | 223 | 1 GLIC_ARATH | Q9fma6 arabidopsis |
| 57 | 7 | 0.8 | 226 | 1 COLI_ONCKE | P10000 o corticotr |
| 58 | 7 | 0.8 | 229 | 1 RIBB_METTH | Q27543 methanobact |
| 59 | 7 | 0.8 | 232 | 1 TRPF_LIPST | Q01128 lipomyces s |
| 60 | 7 | 0.8 | 235 | 1 AURB_CHLAU | P27197 chloroflexu |
| 61 | 7 | 0.8 | 237 | 1 YCY1_SACER | P34698 saccharopol |
| 62 | 7 | 0.8 | 240 | 1 COLJ_ONCMY | Q04618 oncorhynchu |
| 63 | 7 | 0.8 | 254 | 1 GYLR_STRCO | P15360 streptomyce |
| 64 | 7 | 0.8 | 256 | 1 CB4B_LYCES | P27525 lycopersico |
| 65 | 7 | 0.8 | 256 | 1 NORC_VIBCH | Q9x4q5 vibrio chol |
| 66 | 7 | 0.8 | 260 | 1 COX3_XENLA | P00419 xenopus lae |
| 67 | 7 | 0.8 | 260 | 1 REG2_DROME | Q9419 xenopus lae |
| 68 | 7 | 0.8 | 261 | 1 COX3_BRARE | Q9mly4 brachydanio |
| 69 | 7 | 0.8 | 261 | 1 COX3_GADMO | P55777 gadus morhu |
| 70 | 7 | 0.8 | 261 | 1 COX3_POLOR | Q95914 polypterus |
| 71 | 7 | 0.8 | 261 | 1 COX3_SQUAC | Q92248 squaleus aca |
| 72 | 7 | 0.8 | 275 | 1 PHAE_PHAUV | P05088 phaseolus v |
| 73 | 7 | 0.8 | 276 | 1 BACH_HALHP | Q48315 halobacteri |
| 74 | 7 | 0.8 | 276 | 1 BACH_HALHS | Q48314 halobacteri |
| 75 | 7 | 0.8 | 279 | 1 SPBP_RAT | P08723 rattus norv |
| 76 | 7 | 0.8 | 284 | 1 NODO_RHILV | P15728 rhizobium l |
| 77 | 7 | 0.8 | 302 | 1 Y195_VIBCH | Q9kxv1 vibrio chol |
| 78 | 7 | 0.8 | 307 | 1 RUVE_MYCGE | Q49425 mycoplasma |
| 79 | 7 | 0.8 | 307 | 1 RUVE_MYCPN | P75242 mycoplasma |
| 80 | 7 | 0.8 | 308 | 1 ERA_BRAJA | Q69162 bradyrhizob |
| 81 | 7 | 0.8 | 311 | 1 FLIM_BUCAP | Q8ka39 buchnera ap |
| 82 | 7 | 0.8 | 320 | 1 ACC3_CUCME | P54847 cucumis mel |
| 83 | 7 | 0.8 | 320 | 1 F4ST_FLACH | P52837 flaveria ch |
| 84 | 7 | 0.8 | 323 | 1 CD47_HUMAN | Q08722 homo sapien |
| 85 | 7 | 0.8 | 324 | 1 PSC2_MYCTU | Q86344 mycobacteri |
| 86 | 7 | 0.8 | 326 | 1 TGL2_YEAST | P54857 saccharomyc |
| 87 | 7 | 0.8 | 331 | 1 PARG_HUMAN | Q9hb10 homo sapien |
| 88 | 7 | 0.8 | 340 | 1 RECA_RICPR | P41079 rickettsia |
| 89 | 7 | 0.8 | 341 | 1 GPDA_LACLA | Q9cfx6 lactococcus |
| 90 | 7 | 0.8 | 343 | 1 RECA_RICCN | Q92981 rickettsia |
| 91 | 7 | 0.8 | 353 | 1 RRPO_PVSP | P22657 potato viru |
| 92 | 7 | 0.8 | 360 | 1 VG47_BPMU | Q9t1v2 bacterioph |
| 93 | 7 | 0.8 | 363 | 1 OMPC_KLEPN | Q48473 klebsiella |
| 94 | 7 | 0.8 | 366 | 1 AROC_NEIMA | Q9j181 neisseria m |
| 95 | 7 | 0.8 | 366 | 1 AROC_NIMB | Q9j199 neisseria m |
| 96 | 7 | 0.8 | 370 | 1 AROC_CANAL | P79023 candida alb |
| 97 | 7 | 0.8 | 372 | 1 AROC_XYLFA | Q9pd10 xylella fas |
| 98 | 7 | 0.8 | 375 | 1 DNBI_HSVF1 | Q03444 equine herp |
| 99 | 7 | 0.8 | 376 | 1 OPCC_MYCPN | P75553 mycoplasma |
| 100 | 7 | 0.8 | 380 | 1 CYB_LACVV | Q48089 lacerta viv |
| 101 | 7 | 0.8 | 380 | 1 PEL_PSEFL | O59671 pseudomonas |
| 102 | 7 | 0.8 | 380 | 1 TPSA_CABEL | Q77081 caenorhabdi |
| 103 | 7 | 0.8 | 382 | 1 TRPB_DROME | P22809 drosophila |
| 104 | 7 | 0.8 | 406 | 1 TRPB_DROME | P12290 caulobacter |
| 105 | 7 | 0.8 | 414 | 1 GLYA_CAMJE | P24531 campylobact |
| 106 | 7 | 0.8 | 448 | 1 G6PI_FUSNN | Q8rh5 fusobacteri |

| | | | | | | | | | | | | | | |
|-----|-----|---|-----|------|---|-------------|--------------------|-----|---|-----|------|---|-------------|---------------------|
| 107 | 107 | 7 | 0.8 | 449 | 1 | MANB_METJA | Q57842 methanococc | 180 | 7 | 0.8 | 2468 | 1 | MAPB_HUMAN | P46821 homo sapien |
| 108 | 108 | 7 | 0.8 | 454 | 1 | NFM_FIG | P08552 sus scrofa | 181 | 7 | 0.8 | 2649 | 1 | P285_HUMAN | Q9byk8 homo sapien |
| 109 | 109 | 7 | 0.8 | 463 | 1 | POTM_SALTY | P12681 salmonella | 182 | 7 | 0.8 | 3135 | 1 | S230_PLAFO | Q08372 plasmodium |
| 110 | 110 | 7 | 0.8 | 463 | 1 | YCOM_RHORU | P72315 rhodospiril | 183 | 7 | 0.8 | 3430 | 1 | POLG_WNV | P06935 w genome po |
| 111 | 111 | 7 | 0.8 | 465 | 1 | HNFB_HUMAN | Q0ubco homo sapien | 184 | 7 | 0.8 | 3433 | 1 | POLG_KUNJM | P14335 k genome po |
| 112 | 112 | 7 | 0.8 | 465 | 1 | HNFB_MOUSE | O08755 mus musculu | 185 | 7 | 0.8 | 3678 | 1 | DMD_MOUSE | P11531 mus musculu |
| 113 | 113 | 7 | 0.8 | 465 | 1 | HNFB_RAT | P70512 rattus norv | 186 | 7 | 0.8 | 5147 | 1 | PCLO_HUMAN | Q9y6v0 homo sapien |
| 114 | 114 | 7 | 0.8 | 467 | 1 | FLII_BUCAI | P57178 buchnera ap | 187 | 6 | 0.6 | 17 | 1 | LPW_CORGL | P06556 corynebacte |
| 115 | 115 | 7 | 0.8 | 471 | 1 | MURD_STRCO | Q92w9 streptomyce | 188 | 6 | 0.6 | 31 | 1 | PYSG_METBA | P80523 methanosarc |
| 116 | 116 | 7 | 0.8 | 475 | 1 | PRTG_ERWCH | Q07162 erwinia chr | 189 | 6 | 0.6 | 36 | 1 | TLNI_CHICK | P54939 gallus gall |
| 117 | 117 | 7 | 0.8 | 478 | 1 | ASPA_ECOLI | P04422 escherichia | 190 | 6 | 0.6 | 37 | 1 | RL7_CLOFA | P05393 clostridium |
| 118 | 118 | 7 | 0.8 | 479 | 1 | PRTC_ERWCH | P16317 erwinia chr | 191 | 6 | 0.6 | 38 | 1 | RL7_VIBCO | P05395 vibrio cost |
| 119 | 119 | 7 | 0.8 | 483 | 1 | BCA_STRVL | P33569 streptomyce | 192 | 6 | 0.6 | 40 | 1 | UC12_MAIZE | P80618 zea mays (m |
| 120 | 120 | 7 | 0.8 | 486 | 1 | PRZN_SERSP | P07268 serratia sp | 193 | 6 | 0.6 | 42 | 1 | MPMI_SCHPO | P34068 schizosacch |
| 121 | 121 | 7 | 0.8 | 487 | 1 | PRZN_SERWA | P23694 serratia ma | 194 | 6 | 0.6 | 60 | 1 | LHA2_RHOCA | P34068 schizosacch |
| 122 | 122 | 7 | 0.8 | 492 | 1 | SYK_THETH | P41255 thermus the | 195 | 6 | 0.6 | 65 | 1 | OMP_IJOMI | P80045 locusta mig |
| 123 | 123 | 7 | 0.8 | 502 | 1 | VL2_HPV48 | Q80925 human papil | 196 | 6 | 0.6 | 65 | 1 | SAS2_BACCR | P06554 bacillus ce |
| 124 | 124 | 7 | 0.8 | 534 | 1 | FMO2_CAVPO | P36366 cavia porce | 197 | 6 | 0.6 | 71 | 1 | IF1_IJEPIN | Q9xd14 leptospira |
| 125 | 125 | 7 | 0.8 | 537 | 1 | CPA2_MUSDO | O18635 musca domes | 198 | 6 | 0.6 | 71 | 1 | YE34_ANASP | Q44148 anabaena sp |
| 126 | 126 | 7 | 0.8 | 538 | 1 | NRMI_FIG | O77741 sus scrofa | 199 | 6 | 0.6 | 72 | 1 | RPON_THEVO | Q979k0 thermoplas |
| 127 | 127 | 7 | 0.8 | 547 | 1 | GTRI_LEIDO | Q01440 leishmania | 200 | 6 | 0.6 | 72 | 1 | MCHE_ECOLI | Q9rm33 escherichia |
| 128 | 128 | 7 | 0.8 | 557 | 1 | HNFB_HUMAN | P35680 homo sapien | 201 | 6 | 0.6 | 75 | 1 | DMS2_PHYBI | Q902k5 phyllomedusa |
| 129 | 129 | 7 | 0.8 | 557 | 1 | HNFB_RAT | P23899 rattus norv | 202 | 6 | 0.6 | 80 | 1 | DRG2_PACDA | Q93453 pachymedusa |
| 130 | 130 | 7 | 0.8 | 557 | 1 | THS2_HALVO | O30560 halobacteri | 203 | 6 | 0.6 | 81 | 1 | DMS2_PHYBI | P31107 phyllomedusa |
| 131 | 131 | 7 | 0.8 | 558 | 1 | HNFB_MOUSE | P27889 mus musculu | 204 | 6 | 0.6 | 83 | 1 | MXIH_SHIFL | Q06079 shigella fl |
| 132 | 132 | 7 | 0.8 | 559 | 1 | HNFB_FIG | Q33365 sus scrofa | 205 | 6 | 0.6 | 84 | 1 | EX7S_HAEIN | P43914 haemophilus |
| 133 | 133 | 7 | 0.8 | 560 | 1 | TATR_NPVCF | P41716 choristoneu | 206 | 6 | 0.6 | 84 | 1 | SASG_BACSU | P07784 bacillus su |
| 134 | 134 | 7 | 0.8 | 585 | 1 | YEJM_HAEIN | P44898 haemophilus | 207 | 6 | 0.6 | 84 | 1 | UNGI_BPPB2 | P14739 bacterioph |
| 135 | 135 | 7 | 0.8 | 594 | 1 | RMUC_NEIMA | Q9jwg3 neisseria m | 208 | 6 | 0.6 | 84 | 1 | YOAF_ECOLI | P76244 escherichia |
| 136 | 136 | 7 | 0.8 | 594 | 1 | RMUC_NEIMA | Q9jxh2 neisseria m | 209 | 6 | 0.6 | 85 | 1 | RL28_FUSNN | Q8rd9 fusobacteri |
| 137 | 137 | 7 | 0.8 | 617 | 1 | CVG2_HUMAN | O75343 homo sapien | 210 | 6 | 0.6 | 87 | 1 | RPOH_THEVO | Q979f2 thermoplas |
| 138 | 138 | 7 | 0.8 | 622 | 1 | SGA4_DROME | P51905 drosophila | 211 | 6 | 0.6 | 94 | 1 | CH10_THEBR | Q60023 thermoanaer |
| 139 | 139 | 7 | 0.8 | 635 | 1 | SVT_METAC | Q8tlx7 methanosarc | 212 | 6 | 0.6 | 95 | 1 | YD82_ACRFU | Q28899 archaeoglob |
| 140 | 140 | 7 | 0.8 | 637 | 1 | DXS_RHILO | Q985y3 rhizobium l | 213 | 6 | 0.6 | 95 | 1 | YY1_ORVSA | Q23810 oryza sativ |
| 141 | 141 | 7 | 0.8 | 639 | 1 | DXS_AGRTS | Q9uhd7 agrobacteri | 214 | 6 | 0.6 | 96 | 1 | CH10_NEIMB | Q9jxm4 neisseria m |
| 142 | 142 | 7 | 0.8 | 643 | 1 | DXS_BRUME | Q9yfm2 brucella me | 215 | 6 | 0.6 | 96 | 1 | I458_VIBCH | Q9kug7 vibrio chol |
| 143 | 143 | 7 | 0.8 | 643 | 1 | DXS_BRUSU | Q9g292 brucella su | 216 | 6 | 0.6 | 98 | 1 | IM08_SCHPO | Q90783 schizosacch |
| 144 | 144 | 7 | 0.8 | 645 | 1 | DXS_RHIME | Q9zrj1 rhizobium m | 217 | 6 | 0.6 | 100 | 1 | REGN_BPP22 | P04891 bacterioph |
| 145 | 145 | 7 | 0.8 | 646 | 1 | CGLI_CANAL | P24866 candida alb | 218 | 6 | 0.6 | 102 | 1 | ACYM_ACHPL | P14620 anas platyr |
| 146 | 146 | 7 | 0.8 | 646 | 1 | T3MO_BPPI | P08763 bacterioph | 219 | 6 | 0.6 | 102 | 1 | ACYM_CHICK | P07031 gallus gall |
| 147 | 147 | 7 | 0.8 | 654 | 1 | DNAK_CHLMU | P56836 chlamydia m | 220 | 6 | 0.6 | 102 | 1 | ACYM_MELGA | P00821 meleagris g |
| 148 | 148 | 7 | 0.8 | 659 | 1 | DNAK_CHLTR | P17821 chlamydia t | 221 | 6 | 0.6 | 103 | 1 | C552_NITEU | P95329 nitrosomona |
| 149 | 149 | 7 | 0.8 | 670 | 1 | CATA_PENJA | P11934 penicillium | 222 | 6 | 0.6 | 103 | 1 | RL24_BACSU | P12876 bacillus su |
| 150 | 150 | 7 | 0.8 | 685 | 1 | ATKB_CLOAB | O52328 clostridium | 223 | 6 | 0.6 | 103 | 1 | RR6_CVACA | Q19917 cyanidium c |
| 151 | 151 | 7 | 0.8 | 703 | 1 | S261_RAT | P53380 rattus norv | 224 | 6 | 0.6 | 104 | 1 | YGY5_YEAST | P53059 saccharomyc |
| 152 | 152 | 7 | 0.8 | 781 | 1 | COAT_PAVHB | P07299 human parvo | 225 | 6 | 0.6 | 105 | 1 | RL24_THEMEA | P38513 thermotoga |
| 153 | 153 | 7 | 0.8 | 784 | 1 | SVX15_DROME | P40657 drosophila | 226 | 6 | 0.6 | 105 | 1 | Y3C3_STRCO | Q53867 streptomyce |
| 154 | 154 | 7 | 0.8 | 792 | 1 | Y85K_SSV1 | P20210 sulfolobus | 227 | 6 | 0.6 | 106 | 1 | Y0B5_MYCLE | Q49723 mycobacteri |
| 155 | 155 | 7 | 0.8 | 817 | 1 | TGM1_HUMAN | P22735 homo sapien | 228 | 6 | 0.6 | 106 | 1 | YCIH_HAEIN | P45116 haemophilus |
| 156 | 156 | 7 | 0.8 | 827 | 1 | PTF1_RHOCA | P23388 r multiphos | 229 | 6 | 0.6 | 107 | 1 | IHFA_BRUME | Q8ygio brucella me |
| 157 | 157 | 7 | 0.8 | 843 | 1 | NUOG_STRCO | Q9xaro streptomyce | 230 | 6 | 0.6 | 107 | 1 | YI46_SINY3 | P71056 synechocyst |
| 158 | 158 | 7 | 0.8 | 847 | 1 | DNLI_CAEEL | Q27474 caenorhabdi | 231 | 6 | 0.6 | 108 | 1 | HDEB_ECOLI | P26605 escherichia |
| 159 | 159 | 7 | 0.8 | 886 | 1 | RFC1_DROME | P35600 drosophila | 232 | 6 | 0.6 | 109 | 1 | ZM33_MAIZE | Q9abf9 caulobacter |
| 160 | 160 | 7 | 0.8 | 1034 | 1 | ACRF_ECOLI | P24181 escherichia | 233 | 6 | 0.6 | 109 | 1 | HIS2_CLOAB | Q97kh6 clostridium |
| 161 | 161 | 7 | 0.8 | 1041 | 1 | TLR8_HUMAN | Q9nr97 homo sapien | 234 | 6 | 0.6 | 110 | 1 | LOLI_MOUSE | P97873 mus musculu |
| 162 | 162 | 7 | 0.8 | 1138 | 1 | C7AB_BACUT | Q33749 bacillus th | 235 | 6 | 0.6 | 110 | 1 | IHFA_AGRTS | Q8ug61 agrobacteri |
| 163 | 163 | 7 | 0.8 | 1138 | 1 | C7AB_BACUC | Q45707 bacillus th | 236 | 6 | 0.6 | 112 | 1 | IHFA_RHIME | Q92qt3 rhizobium m |
| 164 | 164 | 7 | 0.8 | 1138 | 1 | C7AB_BACUC | P45708 bacillus th | 237 | 6 | 0.6 | 112 | 1 | SMS1_ICTPU | P01171 ictalurus p |
| 165 | 165 | 7 | 0.8 | 1209 | 1 | DNBI_HSVBE | P28932 equine herp | 238 | 6 | 0.6 | 114 | 1 | PT19_STYPL | P28211 styela plic |
| 166 | 166 | 7 | 0.8 | 1287 | 1 | VAC2_HELPY | Q48245 helicobacte | 239 | 6 | 0.6 | 115 | 1 | RL18_MYCCA | P04453 mycoplasma |
| 167 | 167 | 7 | 0.8 | 1288 | 1 | VACA_HELPJ | Q92kw5 helicobacte | 240 | 6 | 0.6 | 116 | 1 | RL20_CANJE | Q9piq0 campylobact |
| 168 | 168 | 7 | 0.8 | 1290 | 1 | BXB_CLOBO | P10844 clostridium | 241 | 6 | 0.6 | 117 | 1 | RL20_AQUAE | Q87086 aquifex aeo |
| 169 | 169 | 7 | 0.8 | 1290 | 1 | VACA_HELPY | P55981 helicobacte | 242 | 6 | 0.6 | 118 | 1 | YE16_HAEIN | P44188 haemophilus |
| 170 | 170 | 7 | 0.8 | 1291 | 1 | VAC4_HELPY | Q48258 helicobacte | 243 | 6 | 0.6 | 118 | 1 | Y15C_ECOLI | P19770 escherichia |
| 171 | 171 | 7 | 0.8 | 1296 | 1 | VAC1_HELPY | Q48247 helicobacte | 244 | 6 | 0.6 | 118 | 1 | NU3C_LUPLU | P52765 lupinus lut |
| 172 | 172 | 7 | 0.8 | 1305 | 1 | RPOC_UREPA | Q9pqv5 ureaplasma | 245 | 6 | 0.6 | 120 | 1 | RL7_ECOLI | P02392 escherichia |
| 173 | 173 | 7 | 0.8 | 1609 | 1 | CTPI_MYCLE | O53114 mycobacteri | 246 | 6 | 0.6 | 120 | 1 | RL7_SALTY | P18081 salmonella |
| 174 | 174 | 7 | 0.8 | 1625 | 1 | CTPI_MYCTU | Q10900 mycobacteri | 247 | 6 | 0.6 | 120 | 1 | RL7_CLOPE | Q8xh7 clostridium |
| 175 | 175 | 7 | 0.8 | 2300 | 1 | CYAA_NEUCR | Q01631 neurospora | 248 | 6 | 0.6 | 121 | 1 | RL7_UREPA | Q9prd4 ureaplasma |
| 176 | 176 | 7 | 0.8 | 2314 | 1 | AKAG_RAT | Q9wvc7 rattus norv | 249 | 6 | 0.6 | 121 | 1 | RL7_VIBCH | Q9rdv1 vibrio chol |
| 177 | 177 | 7 | 0.8 | 2319 | 1 | AKAG_HUMAN | Q13023 homo sapien | 250 | 6 | 0.6 | 121 | 1 | RL7_XANAC | Q8pnc1 xanthomonas |
| 178 | 178 | 7 | 0.8 | 2459 | 1 | MAPB_RAT | P15205 rattus norv | 251 | 6 | 0.6 | 121 | 1 | RL7_XANCP | Q8pc57 xanthomonas |
| 179 | 179 | 7 | 0.8 | 2464 | 1 | MAPB_MOUSE | P14873 mus musculu | 252 | 6 | 0.6 | 121 | 1 | RL7_XANCP | Q8pc57 xanthomonas |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|---|------------|--------------------|-----|---|-----|-----|---|------------|---------------------|
| 253 | 6 | 0.6 | 122 | 1 | RL7_HAEIN | P44348 haemophilus | 326 | 6 | 0.6 | 141 | 1 | HBA_COTJA | P24589 coturnix co |
| 254 | 6 | 0.6 | 122 | 1 | RL7_HALEU | P07472 halophilic | 327 | 6 | 0.6 | 141 | 1 | HBA_CYGOL | P01992 cygnus olor |
| 255 | 6 | 0.6 | 122 | 1 | RL7_HALPR | P41134 haloanaerob | 328 | 6 | 0.6 | 141 | 1 | HBA_LARRI | P08260 laryus ridib |
| 256 | 6 | 0.6 | 122 | 1 | RL7_PASMU | Q9ck90 pasteurella | 329 | 6 | 0.6 | 141 | 1 | HBA_PASMO | P07407 passer mont |
| 257 | 6 | 0.6 | 122 | 1 | RL7_VIBPA | Q87kq3 vibrio para | 330 | 6 | 0.6 | 141 | 1 | HBA_PHACA | P10780 phalacrocor |
| 258 | 6 | 0.6 | 122 | 1 | RL7_XYLFa | Q9pa85 xyella fas | 331 | 6 | 0.6 | 141 | 1 | HBA_RAT | P01946 rattus norv |
| 259 | 6 | 0.6 | 122 | 1 | YMI9_YEAST | Q04502 saccharomyc | 332 | 6 | 0.6 | 141 | 1 | HBA_STUVU | P01997 sturnus vul |
| 260 | 6 | 0.6 | 122 | 1 | RL7_CLOAB | Q97e98 clostridium | 333 | 6 | 0.6 | 141 | 1 | HBA_TURME | P14522 turdus meru |
| 261 | 6 | 0.6 | 123 | 1 | VST1_HEVBU | Q29e89 hepatitis e | 334 | 6 | 0.6 | 141 | 1 | PRO1_STRPU | P32006 strongyloce |
| 262 | 6 | 0.6 | 123 | 1 | VST1_HEVME | Q03499 hepatitis e | 335 | 6 | 0.6 | 141 | 1 | RL23_TRYCR | Q94776 trypanosoma |
| 263 | 6 | 0.6 | 123 | 1 | VST1_HEVYU | Q04612 hepatitis e | 336 | 6 | 0.6 | 141 | 1 | YQBS_BACSU | P45934 bacillus su |
| 264 | 6 | 0.6 | 124 | 1 | ACPS_CLOAB | Q971r5 clostridium | 337 | 6 | 0.6 | 142 | 1 | YQAA_ECOLI | P76631 escherichia |
| 265 | 6 | 0.6 | 124 | 1 | R22B_ARATH | Q9m9w1 arabidopsis | 338 | 6 | 0.6 | 143 | 1 | VG32_VZVD | P09285 varicella-2 |
| 266 | 6 | 0.6 | 124 | 1 | R22C_ARATH | Q9fe58 arabidopsis | 339 | 6 | 0.6 | 143 | 1 | YIBN_ECOLI | P37688 escherichia |
| 267 | 6 | 0.6 | 124 | 1 | RL7_BRUNE | P41106 brucella me | 340 | 6 | 0.6 | 144 | 1 | FLAV_TRIER | O52659 trichodesmi |
| 268 | 6 | 0.6 | 125 | 1 | C556_AGRTA | P00141 agrobacteri | 341 | 6 | 0.6 | 145 | 1 | PA21_BOVIN | P00593 bos taurus |
| 269 | 6 | 0.6 | 125 | 1 | RL7_AGRT5 | Q8ue07 agrobacteri | 342 | 6 | 0.6 | 146 | 1 | RPBB_YEAST | P20436 saccharomyc |
| 270 | 6 | 0.6 | 125 | 1 | RL7_LIBAF | P41189 liberibacte | 343 | 6 | 0.6 | 146 | 1 | YRFB_ECOLI | P45751 escherichia |
| 271 | 6 | 0.6 | 125 | 1 | RL7_RHILO | Q98n67 rhizobium 1 | 344 | 6 | 0.6 | 147 | 1 | YA37_MYCPN | P75077 mycoplasma |
| 272 | 6 | 0.6 | 125 | 1 | RL7_RICCN | Q92j87 rickettsia | 345 | 6 | 0.6 | 148 | 1 | CTR6_SCHPO | Q9usv7 schizosacch |
| 273 | 6 | 0.6 | 125 | 1 | RL7_RICPR | Q9ze21 rickettsia | 346 | 6 | 0.6 | 148 | 1 | PSAL_SYNEH | P25902 synchococc |
| 274 | 6 | 0.6 | 125 | 1 | RL7_THETN | Q8r7u5 thermoanaer | 347 | 6 | 0.6 | 149 | 1 | IL2_HORSE | P37997 equus cabal |
| 275 | 6 | 0.6 | 125 | 1 | YOM2_PHOPR | P29740 photobacter | 348 | 6 | 0.6 | 149 | 1 | MLE_BRAFL | Q17133 branchiosto |
| 276 | 6 | 0.6 | 126 | 1 | RL7_DESYM | P02393 desulfovibr | 349 | 6 | 0.6 | 149 | 1 | YB29_MYCPN | P75346 mycoplasma |
| 277 | 6 | 0.6 | 126 | 1 | RL7_RHIME | Q92qh8 rhizobium m | 350 | 6 | 0.6 | 150 | 1 | EXBB_HAEDU | O51808 haemophilus |
| 278 | 6 | 0.6 | 127 | 1 | KRC1_CHICK | P25692 gallus gall | 351 | 6 | 0.6 | 150 | 1 | MLE_DICDI | P09402 dictyosteli |
| 279 | 6 | 0.6 | 127 | 1 | R22A_ARATH | Q9erx7 arabidopsis | 352 | 6 | 0.6 | 150 | 1 | NIFW_RHISN | P55689 rhizobium s |
| 280 | 6 | 0.6 | 127 | 1 | RL7_CAUCR | Q9aa07 caulobacter | 353 | 6 | 0.6 | 150 | 1 | PA2C_MOUSE | P48076 mus musculu |
| 281 | 6 | 0.6 | 127 | 1 | RL7_SYNY3 | P23349 synchocyst | 354 | 6 | 0.6 | 150 | 1 | PSAL_ODOSI | P39878 rattus norv |
| 282 | 6 | 0.6 | 127 | 1 | RT08_ACACA | P46757 acanthamoeb | 355 | 6 | 0.6 | 150 | 1 | PA2C_RAT | P49486 otomella s |
| 283 | 6 | 0.6 | 128 | 1 | GC5H_YERPE | Q8zh17 yersinia pe | 356 | 6 | 0.6 | 151 | 1 | Y22B_AQUAE | O66419 aquifex aeo |
| 284 | 6 | 0.6 | 128 | 1 | KDGL_HELPJ | Q9zle0 helicobacte | 357 | 6 | 0.6 | 152 | 1 | EXBB_PASHA | P72202 pasteurella |
| 285 | 6 | 0.6 | 128 | 1 | KDGL_HELPY | P56411 helicobacte | 358 | 6 | 0.6 | 152 | 1 | IAA2_HORVU | P13691 hordeum vul |
| 286 | 6 | 0.6 | 128 | 1 | RL7_AQUAE | O67761 aquifex aeo | 359 | 6 | 0.6 | 152 | 1 | PYRI_THEVO | Q97b28 thermoplasm |
| 287 | 6 | 0.6 | 129 | 1 | LV1D_MOUSE | P01726 mus musculu | 360 | 6 | 0.6 | 153 | 1 | IL2_HUMAN | F01585 homo sapien |
| 288 | 6 | 0.6 | 129 | 1 | MFA4_BOVIN | P55918 bos taurus | 361 | 6 | 0.6 | 153 | 1 | YIBG_ECOLI | P32106 escherichia |
| 289 | 6 | 0.6 | 129 | 1 | RL7_ANASP | Q8ylj5 anabaena sp | 362 | 6 | 0.6 | 154 | 1 | IL2_CERTO | P46649 cercocerib |
| 290 | 6 | 0.6 | 129 | 1 | TDCP_ECOLI | P42611 escherichia | 363 | 6 | 0.6 | 154 | 1 | IL2_FELCA | Q07885 felis silve |
| 291 | 6 | 0.6 | 129 | 1 | YB58_METJA | Q58557 methanococc | 364 | 6 | 0.6 | 154 | 1 | IL2_MACFA | Q29615 macaca fasc |
| 292 | 6 | 0.6 | 130 | 1 | F14A_HUMAN | Q9h2x8 homo sapien | 365 | 6 | 0.6 | 154 | 1 | IL2_MACMU | P51498 macaca mula |
| 293 | 6 | 0.6 | 130 | 1 | RL14_LEPIN | Q9xd26 leptospira | 366 | 6 | 0.6 | 154 | 1 | IL2_MIRAN | O62641 mirounga an |
| 294 | 6 | 0.6 | 130 | 1 | Y122_BUCAI | P57222 bucherna ap | 367 | 6 | 0.6 | 154 | 1 | YQAR_BACSU | P45914 bacillus su |
| 295 | 6 | 0.6 | 130 | 1 | Y483_DEIRA | Q9rx32 deinococcus | 368 | 6 | 0.6 | 155 | 1 | Y115_METJA | O57579 methanococc |
| 296 | 6 | 0.6 | 133 | 1 | GOP_BPP4 | P13058 bacterioph | 369 | 6 | 0.6 | 156 | 1 | ATPF_VIBAL | P12989 vibrio algi |
| 297 | 6 | 0.6 | 134 | 1 | H2AV_YEAST | Q12692 saccharomyc | 370 | 6 | 0.6 | 156 | 1 | PMAD_BACNO | P13253 bacteroides |
| 298 | 6 | 0.6 | 134 | 1 | Y886_HAEIN | P44069 haemophilus | 371 | 6 | 0.6 | 156 | 1 | VG55_BMWLS | Q05285 mycobacteri |
| 299 | 6 | 0.6 | 135 | 1 | LGUL_HAEIN | P44638 haemophilus | 372 | 6 | 0.6 | 157 | 1 | EGD1_YEAST | Q02642 saccharomyc |
| 300 | 6 | 0.6 | 136 | 1 | YGHG_ECOLI | Q46835 escherichia | 373 | 6 | 0.6 | 157 | 1 | FKB1_METJA | Q57726 methanococc |
| 301 | 6 | 0.6 | 137 | 1 | GTH2_ACALA | Q90225 acanthopagr | 374 | 6 | 0.6 | 157 | 1 | NIFU_AQUAE | O67045 aquifex aeo |
| 302 | 6 | 0.6 | 137 | 1 | RL19_NEPOL | Q8tkz4 nephroselmi | 375 | 6 | 0.6 | 157 | 1 | PR1_MEDSA | Q43560 medicago sa |
| 303 | 6 | 0.6 | 138 | 1 | CXA4_PIG | Q29559 sus scrofa | 376 | 6 | 0.6 | 157 | 1 | SSSF_PLAFO | P17503 plasmodium |
| 304 | 6 | 0.6 | 139 | 1 | YOGF_HAEIN | P43981 haemophilus | 377 | 6 | 0.6 | 157 | 1 | YGL7_PASMU | Q9ckj8 pasteurella |
| 305 | 6 | 0.6 | 140 | 1 | HBB_RANES | P02134 rana escul | 378 | 6 | 0.6 | 158 | 1 | RRJ3_PEA | O22386 oryza sativ |
| 306 | 6 | 0.6 | 140 | 1 | POP7_HUMAN | Q75817 homo sapien | 379 | 6 | 0.6 | 158 | 1 | RL12_ORISA | F14710 pisum sativ |
| 307 | 6 | 0.6 | 141 | 1 | HBA1_CATWA | P82111 catharacta | 380 | 6 | 0.6 | 159 | 1 | RL22_MYCPN | P75575 mycoplasma |
| 308 | 6 | 0.6 | 141 | 1 | HBA1_IGUIG | P18974 iguana igua | 381 | 6 | 0.6 | 159 | 1 | YIB6_YEAST | P40548 saccharomyc |
| 309 | 6 | 0.6 | 141 | 1 | HBA2_BURBU | Q8tsn8 bubalus bub | 382 | 6 | 0.6 | 160 | 1 | BCHF_RHOSH | O53222 rhodobacter |
| 310 | 6 | 0.6 | 141 | 1 | HBA2_ANAPL | P04442 anas platyr | 383 | 6 | 0.6 | 160 | 1 | PHAA_CYAPA | P00316 cyanophora |
| 311 | 6 | 0.6 | 141 | 1 | HBA2_ANANP | P04238 anser anser | 384 | 6 | 0.6 | 161 | 1 | 7B2_XENLA | F18844 xenopus lae |
| 312 | 6 | 0.6 | 141 | 1 | HBA2_ANGIN | P04239 anser indic | 385 | 6 | 0.6 | 161 | 1 | PHAB_AGLNE | P28556 aglaothamni |
| 313 | 6 | 0.6 | 141 | 1 | HBA2_BRACA | P04240 branta cana | 386 | 6 | 0.6 | 162 | 1 | DYR_LACCA | P00381 lactobacill |
| 314 | 6 | 0.6 | 141 | 1 | HBA2_CAIMO | P02003 cairina mos | 387 | 6 | 0.6 | 162 | 1 | MOTB_BPT4 | Q01437 bacterioph |
| 315 | 6 | 0.6 | 141 | 1 | HBA2_CHLME | P07035 chloephaga | 388 | 6 | 0.6 | 163 | 1 | ATPX_OCHNE | O40608 ochtrophae |
| 316 | 6 | 0.6 | 141 | 1 | HBA2_ANAPE | P22740 anas penelo | 389 | 6 | 0.6 | 163 | 1 | GC5H_MESCR | P93255 mesembryant |
| 317 | 6 | 0.6 | 141 | 1 | HBA2_ANAPL | P01986 anas platyr | 390 | 6 | 0.6 | 163 | 1 | MLC2_YEAST | Q06580 saccharomyc |
| 318 | 6 | 0.6 | 141 | 1 | HBA2_ANAPP | P01988 anas platyr | 391 | 6 | 0.6 | 164 | 1 | YAF9_METAC | P58828 methanosarc |
| 319 | 6 | 0.6 | 141 | 1 | HBA2_ANANP | P01989 anser anser | 392 | 6 | 0.6 | 165 | 1 | OGT_BACSU | P11742 bacillus su |
| 320 | 6 | 0.6 | 141 | 1 | HBA2_ANNSE | P01985 anseranas s | 393 | 6 | 0.6 | 166 | 1 | PMPA_CANBO | P14292 candida boi |
| 321 | 6 | 0.6 | 141 | 1 | HBA2_APTFO | P01980 aptenodyces | 394 | 6 | 0.6 | 166 | 1 | PMPB_CANBO | P14293 candida boi |
| 322 | 6 | 0.6 | 141 | 1 | HBA2_ARAAR | P01996 ara araraun | 395 | 6 | 0.6 | 166 | 1 | RL10_PSEAM | Q9hw7 pseudomonas |
| 323 | 6 | 0.6 | 141 | 1 | HBA2_BRACA | P01991 branta cana | 396 | 6 | 0.6 | 166 | 1 | RS5_STAAM | Q99838 staphylococ |
| 324 | 6 | 0.6 | 141 | 1 | HBA2_CAIMO | P01987 cairina mos | 397 | 6 | 0.6 | 167 | 1 | DYR_ENTFC | P00380 enterococcu |
| 325 | 6 | 0.6 | 141 | 1 | HBA2_CHLME | P07034 chloephaga | 398 | 6 | 0.6 | 167 | 1 | RS5_SHEON | P59124 shewanella |

| | | | | | | | | | | | |
|-----|---|------|-------|---------|--------------|-----|---|------|-------|---------|--------------|
| 399 | 1 | YP03 | BORBU | Q44849 | borrelia bu | 472 | 1 | CD8B | HUMAN | P10966 | homo sapien |
| 400 | 1 | LED1 | METKA | Q8tx94 | methanopyru | 473 | 1 | CD8B | PONPY | P30434 | pongo pygma |
| 401 | 1 | CFP7 | CAEEL | P2015 | caenorhabdi | 474 | 1 | KTHY | PYRFU | Q8u071 | pyrococcus |
| 402 | 1 | RL10 | CORGL | Q8nt29 | corynebacte | 475 | 1 | HET3 | RADMG | Q9u6x1 | radianthus |
| 403 | 1 | CYP1 | ARATH | P34790 | arabidopsis | 476 | 1 | ROPB | RHLV | Q52866 | thizobium 1 |
| 404 | 1 | CYP2 | CAEEL | P52010 | caenorhabdi | 477 | 1 | SDF2 | HUMAN | Q99470 | homo sapien |
| 405 | 1 | CYP3 | CAEEL | P52011 | caenorhabdi | 478 | 1 | SDF2 | MOUSE | Q9dct5 | mus musculus |
| 406 | 1 | YB06 | VACCV | P21002 | vaccinia vi | 479 | 1 | YH73 | METTH | Q27801 | methanobact |
| 407 | 1 | Y096 | NFVAC | P41656 | autographa | 480 | 1 | IL6 | CERTO | P46650 | cercocebus |
| 408 | 1 | AX15 | ARATH | P33078 | arabidopsis | 481 | 1 | IL6 | HUMAN | P05231 | homo sapien |
| 409 | 1 | HSCB | HAEIN | Q57006 | haemophilus | 482 | 1 | IL6 | MACFA | P79341 | macaca fasc |
| 410 | 1 | CVT2 | STOHE | P07845 | stoichactis | 483 | 1 | IL6 | MACMU | P51494 | macaca mula |
| 411 | 1 | ET2 | MOUSE | P22389 | mus musculus | 484 | 1 | PCAJ | PSEPK | Q01104 | pseudomonas |
| 412 | 1 | NU6M | CANPA | Q9zz56 | canis famil | 485 | 1 | RPSD | MYCTU | Q50712 | mycobacteri |
| 413 | 1 | YF49 | MYCTU | Q10777 | mycobacteri | 486 | 1 | HIS5 | LEGPN | Q3rd31 | legionella |
| 414 | 1 | CYT1 | STOHE | P81662 | stoichactis | 487 | 1 | LEXA | XANAC | Q86050 | xanthomonas |
| 415 | 1 | PACA | HUMAN | P18509 | h pituitary | 488 | 1 | LEXA | XANCA | Q93mq9 | xanthomonas |
| 416 | 1 | ARGR | STRCL | P95721 | streptomyce | 489 | 1 | LEXA | XANCP | Q8p9x2 | xanthomonas |
| 417 | 1 | ATPD | STREO | O50156 | streptococc | 490 | 1 | PURQ | THEMA | Q9x0x2 | thermotoga |
| 418 | 1 | ARGR | STRCO | Q911as | streptomyce | 491 | 1 | PYRQ | HAEIN | P43855 | haemophilus |
| 419 | 1 | Y053 | SYNY3 | Q55142 | synechocyst | 492 | 1 | Y767 | MYCTU | P71822 | mycobacteri |
| 420 | 1 | PORC | METTH | O27772 | methanobact | 493 | 1 | YE54 | HAEIN | P44202 | haemophilus |
| 421 | 1 | IF3 | RHOSH | O33567 | rhodobacter | 494 | 1 | YJH2 | YEAST | P40359 | saccharomyc |
| 422 | 1 | Y06E | BPT4 | P13311 | bacterioph | 495 | 1 | GIDB | PSEAE | Q9htc10 | pseudomonas |
| 423 | 1 | APT | CAMJE | Q9pp06 | campylobact | 496 | 1 | Y315 | TREPA | P56822 | treponema p |
| 424 | 1 | RELX | PIG | P01348 | sus scrofa | 497 | 1 | PSBP | BRATU | Q96334 | brassica ju |
| 425 | 1 | YGL1 | ECOLI | P39834 | escherichia | 498 | 1 | SWSH | DROME | Q9vj36 | drosophila |
| 426 | 1 | RK5 | MESVI | Q9muu5 | mesostigma | 499 | 1 | NEP1 | SULTO | Q96yp4 | suifolobus |
| 427 | 1 | YCF4 | CHAGL | Q8m9x4 | chaetosphae | 500 | 1 | RAS | SCHPO | P08647 | schizosacch |
| 428 | 1 | YF3D | DROME | Q9vgp6 | drorophila | 501 | 1 | RT02 | ARATH | Q9gcb9 | arabidopsis |
| 429 | 1 | RFP | VIBPA | Q8rfs5 | vibrio para | 502 | 1 | YGH8 | SALTY | P18951 | salmonella |
| 430 | 1 | RP06 | FOPEV | Q85280 | fowlpox vir | 503 | 1 | CLD6 | HUMAN | P56747 | homo sapien |
| 431 | 1 | YB15 | ARCFU | O29150 | archaeoglob | 504 | 1 | RADB | PYRKO | P95547 | pyrococcus |
| 432 | 1 | KAD | XANAC | Q9phn23 | xanthomonas | 505 | 1 | GTAT | CHICK | Q08392 | gallus gall |
| 433 | 1 | HPK1 | MYCTU | O06276 | mycobacteri | 506 | 1 | SDFL | HUMAN | Q9hcn8 | homo sapien |
| 434 | 1 | KTHY | METJA | Q57491 | methanococc | 507 | 1 | SDFL | MOUSE | Q9esp1 | mus musculus |
| 435 | 1 | UBCB | YEAST | P52491 | saccharomyc | 508 | 1 | TCPP | VIBCH | P29485 | vibrio chol |
| 436 | 1 | NUXM | NEUCR | Q02854 | neurospora | 509 | 1 | COLI | CYPCA | Q9y9k4 | c corticotr |
| 437 | 1 | RP2 | MOUSE | Q9wup0 | mus musculus | 510 | 1 | KCV | AQJAE | O67907 | aquifex aeo |
| 438 | 1 | Y573 | METJA | Q57993 | methanococc | 511 | 1 | SODM | HORSE | Q9x841 | equus cabal |
| 439 | 1 | KADA | METKA | Q8tzb0 | methanopyru | 512 | 1 | PLLV | RAT | P34207 | rattus norv |
| 440 | 1 | KIPI | HUMAN | Q99828 | homo sapien | 513 | 1 | RPIA | STRA3 | Q8e4y6 | streptococc |
| 441 | 1 | KIPI | MOUSE | Q920f4 | mus musculus | 514 | 1 | MTGA | STRAS | Q8dz66 | streptococc |
| 442 | 1 | KIPI | RAT | Q920f4 | mus musculus | 515 | 1 | MTGA | ACICA | O24849 | acinotobact |
| 443 | 1 | CFP1 | CAEEL | P52009 | caenorhabdi | 516 | 1 | MYOG | MOUSE | P12979 | mus musculus |
| 444 | 1 | Y210 | AQJAE | O51880 | buchnera ap | 517 | 1 | PURQ | HALN1 | Q9hnu2 | halobacteri |
| 445 | 1 | Y210 | AQJAE | O66404 | aquifex aeo | 518 | 1 | MODB | AZQV1 | P37731 | azotobacter |
| 446 | 1 | KAD | SYNEL | Q8gml4 | synechococc | 519 | 1 | KTHY | XANAC | Q8pf97 | xanthomonas |
| 447 | 1 | TRP6 | CYACA | O19914 | cyanidium c | 520 | 1 | CLDA | HUMAN | P78369 | homo sapien |
| 448 | 1 | PINO | ECOLI | P77170 | escherichia | 521 | 1 | ISTB | BURCE | P55923 | burkholderi |
| 449 | 1 | PINR | ECOLI | P77574 | escherichia | 522 | 1 | PYRH | THEAC | Q9hkv7 | thermoplasm |
| 450 | 1 | RNH2 | AQJAE | O67768 | aquifex aeo | 523 | 1 | Y4MB | RHISN | P55561 | thizobium s |
| 451 | 1 | R85 | DEIRA | Q9r8l1 | deinococcus | 524 | 1 | Y014 | ARCFU | O66958 | aquifex aeo |
| 452 | 1 | RACD | ENTHI | Q24817 | entamoeba h | 525 | 1 | Y756 | AQJAE | Q30257 | archaeoglob |
| 453 | 1 | YB65 | MYCPN | P75318 | mycoplasma | 526 | 1 | C552 | THIFE | P74917 | thiobacillu |
| 454 | 1 | COAE | BRUME | Q8ye21 | brucella me | 527 | 1 | YGEA | ECOLI | P01813 | escherichia |
| 455 | 1 | R8FA | ARATH | O65693 | arabidopsis | 528 | 1 | CLDA | MOUSE | Q920s6 | mus musculus |
| 456 | 1 | PS8 | METJA | P27058 | lycoperisicc | 529 | 1 | CLP2 | ANASP | Q8yq88 | anabaena sp |
| 457 | 1 | RECR | ECOLI | Q58609 | methanococc | 530 | 1 | UPP | PYRAB | Q9u0k1 | pyrococcus |
| 458 | 1 | RECR | SALTY | P12727 | escherichia | 531 | 1 | UPP | PYRFU | Q8ul97 | pyrococcus |
| 459 | 1 | RECR | SALTY | Q8xg25 | salmonella | 532 | 1 | MTGA | NEIGO | Q51005 | neisseria g |
| 460 | 1 | RUVA | CHLTE | Q8kfr0 | chlorobium | 533 | 1 | MTGA | NEIMA | O52423 | neisseria m |
| 461 | 1 | C08G | RABIT | Q28679 | oryctolagus | 534 | 1 | VRPP | LAMB | P03689 | bacterioph |
| 462 | 1 | GDIR | YEAST | Q12434 | saccharomyc | 535 | 1 | Y081 | NFVAC | Q06694 | autographa |
| 463 | 1 | MSX3 | MOUSE | P70354 | mus musculus | 536 | 1 | YX02 | CAEEL | Q11109 | caenorhabdi |
| 464 | 1 | YA57 | ACTAC | O52728 | actinobacil | 537 | 1 | TFE2 | ALCEU | P94136 | alcaligenes |
| 465 | 1 | 3MCH | CLOPE | Q8xha9 | clostridium | 538 | 1 | Y381 | MYCPN | P75219 | mycoplasma |
| 466 | 1 | RUVA | AGRTS | Q8u9k5 | agrobacteri | 539 | 1 | Y680 | RHLIO | Q98m91 | thizobium 1 |
| 467 | 1 | YD06 | NGRT5 | Q8uful | agrobacteri | 540 | 1 | PCNA | SOYBN | P22177 | glycine max |
| 468 | 1 | RR4 | ODOSI | P49492 | odontella s | 541 | 1 | NOLA | BRASN | P50330 | bradyrhizob |
| 469 | 1 | CYPB | CHICK | P24367 | gallus gall | 542 | 1 | PHOU | ZYMMO | Q9x5el | zyimonas m |
| 470 | 1 | YJY2 | YEAST | P47087 | saccharomyc | 543 | 1 | PYRF | FUSN3 | Q8rg83 | fusobacteri |
| 471 | 1 | CD8B | FELCA | P79336 | felis silve | 544 | 1 | SFSA | SYNY3 | P73664 | synechocyst |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|---|------------|--------------------|-----|---|-----|-----|---|------------|---------------------|
| 691 | 6 | 0.6 | 279 | 1 | NIFH_METJA | Q58289 methanococ | 764 | 6 | 0.6 | 305 | 1 | GP7D_CHLMU | O46435 chlamydia m |
| 692 | 6 | 0.6 | 279 | 1 | YB49_STRP3 | Q8k6c3 streptococ | 765 | 6 | 0.6 | 305 | 1 | GP7D_CHLTR | P10561 chlamydia t |
| 693 | 6 | 0.6 | 280 | 1 | YE93_STRP3 | Q99YV0 streptococ | 766 | 6 | 0.6 | 305 | 1 | PYRB_SERMA | P19910 serratia ma |
| 694 | 6 | 0.6 | 279 | 1 | ISPE_OCEIH | Q8eu37 oceanobacil | 767 | 6 | 0.6 | 305 | 1 | RS1B_SYN3 | P74142 synechocyst |
| 695 | 6 | 0.6 | 280 | 1 | NUDC_DEIRA | Q9rv62 deinococcus | 768 | 6 | 0.6 | 305 | 1 | SGAU_MYCPN | P75294 mycoplasma |
| 696 | 6 | 0.6 | 280 | 1 | PANC_XANNC | Q8p110 xanthomonas | 769 | 6 | 0.6 | 305 | 1 | YYAM_BACSU | P37511 bacillus su |
| 697 | 6 | 0.6 | 280 | 1 | PURU_ECOL6 | Q8p110 xanthomonas | 770 | 6 | 0.6 | 306 | 1 | ARGI_BRUME | O59174 brucella me |
| 698 | 6 | 0.6 | 280 | 1 | PURU_ECOL6 | P38480 escherichia | 771 | 6 | 0.6 | 306 | 1 | UCPI_MOUSE | P12242 mus musculus |
| 699 | 6 | 0.6 | 280 | 1 | PURU_ECOL1 | P37051 escherichia | 772 | 6 | 0.6 | 306 | 1 | UCPI_RAT | P04633 rattus norv |
| 700 | 6 | 0.6 | 282 | 1 | Y017_MYCTU | P71726 mycobacteri | 773 | 6 | 0.6 | 307 | 1 | CIW7_HUMAN | Q9Y242 homo sapien |
| 701 | 6 | 0.6 | 282 | 1 | EFTS_CHLPN | Q9h117 thermoplas | 774 | 6 | 0.6 | 307 | 1 | GP7D_CHLPS | O46257 chlamydia p |
| 702 | 6 | 0.6 | 282 | 1 | PHD2_THEAC | Q9wxx9 thermotoga | 775 | 6 | 0.6 | 307 | 1 | PRB2_CORGL | O8ns12 corynebacte |
| 703 | 6 | 0.6 | 283 | 1 | CGIC_CHICK | P55168 gallus gall | 776 | 6 | 0.6 | 307 | 1 | YH33_YEAST | P38747 saccharomyc |
| 704 | 6 | 0.6 | 283 | 1 | DAPF_NEIMA | Q9jv69 neisseria m | 777 | 6 | 0.6 | 308 | 1 | ACPI_ENTHI | P36184 entamoeba h |
| 705 | 6 | 0.6 | 283 | 1 | DAPF_NEIMB | Q9K060 neisseria m | 778 | 6 | 0.6 | 308 | 1 | AEPI_DROME | P39413 drosophila |
| 706 | 6 | 0.6 | 283 | 1 | YQVW_CABEL | Q17439 caenorhabdi | 779 | 6 | 0.6 | 308 | 1 | CYB_AMEBA | P29632 amebophis |
| 707 | 6 | 0.6 | 284 | 1 | FOCA_HAEIN | P43756 haemophilus | 780 | 6 | 0.6 | 308 | 1 | CYB_AMPST | P29633 ampeleon st |
| 708 | 6 | 0.6 | 284 | 1 | PANC_XANCP | Q9p989 xanthomonas | 781 | 6 | 0.6 | 308 | 1 | CYB_COLRU | P29635 colaptes ru |
| 709 | 6 | 0.6 | 285 | 1 | AROE_CAUCR | Q8ac57 caulobacter | 782 | 6 | 0.6 | 308 | 1 | CYB_PARIN | P29638 parus inorn |
| 710 | 6 | 0.6 | 285 | 1 | ARSB_MOUSE | P50429 mus musculu | 783 | 6 | 0.6 | 308 | 1 | CYB_PTPL | P29640 ptilopora |
| 711 | 6 | 0.6 | 285 | 1 | PANC_CAUCR | Q8ac68 caulobacter | 784 | 6 | 0.6 | 308 | 1 | CYB_SCYMA | P29641 scytalopus |
| 712 | 6 | 0.6 | 286 | 1 | PNK1_CAMJE | Q9phm6 campylobact | 785 | 6 | 0.6 | 308 | 1 | CYB_THRDO | P29642 thripophaga |
| 713 | 6 | 0.6 | 286 | 1 | TRT2_SHEEP | P50751 ovis aries | 786 | 6 | 0.6 | 309 | 1 | FABH_AQUAE | O67185 aquifex aeo |
| 714 | 6 | 0.6 | 286 | 1 | YDBC_ECOLI | P25906 escherichia | 787 | 6 | 0.6 | 309 | 1 | NTCB_SYNP7 | P52683 synechococ |
| 715 | 6 | 0.6 | 287 | 1 | FOLD_SALTI | Q50006 salmonella | 788 | 6 | 0.6 | 309 | 1 | PYRB_VIBPA | Q871f8 vibrio para |
| 716 | 6 | 0.6 | 287 | 1 | FOLD_SALTI | P58688 salmonella | 789 | 6 | 0.6 | 309 | 1 | PYRB_VIBS2 | P96174 vibrio sp. |
| 717 | 6 | 0.6 | 287 | 1 | HBD_BACSU | P45856 bacillus su | 790 | 6 | 0.6 | 309 | 1 | PYRB_VIBVU | Q8dcf6 vibrio vuln |
| 718 | 6 | 0.6 | 287 | 1 | TVSY_MYCCE | P47469 mycoplasma | 791 | 6 | 0.6 | 310 | 1 | PYRB_ECOLI | P00479 escherichia |
| 719 | 6 | 0.6 | 288 | 1 | CD80_HUMAN | P33681 homo sapien | 792 | 6 | 0.6 | 310 | 1 | PYRB_SALTY | P08420 salmonella |
| 720 | 6 | 0.6 | 288 | 1 | OTX2_XENLA | Q18133 xenopus lae | 793 | 6 | 0.6 | 310 | 1 | RPOA_CVACA | Q9t1v2 cyanidium c |
| 721 | 6 | 0.6 | 288 | 1 | TRA6_NEIMB | Q00840 neisseria m | 794 | 6 | 0.6 | 310 | 1 | YCVJ_ECOLI | P76049 escherichia |
| 722 | 6 | 0.6 | 289 | 1 | OTX2_ERARE | Q91981 brachydanio | 795 | 6 | 0.6 | 311 | 1 | P3ST_FLACH | P52836 flaveria ch |
| 723 | 6 | 0.6 | 289 | 1 | OTX2_HUMAN | P32243 homo sapien | 796 | 6 | 0.6 | 311 | 1 | PYRB_YERPE | Q82b39 yersinia pe |
| 724 | 6 | 0.6 | 289 | 1 | OTX2_MOUSE | P80206 mus musculu | 797 | 6 | 0.6 | 311 | 1 | YXXF_BACSU | Q07835 bacillus su |
| 725 | 6 | 0.6 | 290 | 1 | BLA1_XANNA | P52700 xanthomonas | 798 | 6 | 0.6 | 312 | 1 | F3ST_FLABI | P52835 flaveria bi |
| 726 | 6 | 0.6 | 290 | 1 | YC38_CYAPA | P48278 cyanophora | 799 | 6 | 0.6 | 312 | 1 | FABD_HAEIN | P43712 haemophilus |
| 727 | 6 | 0.6 | 291 | 1 | ISPA_MICLU | O66126 micrococcus | 800 | 6 | 0.6 | 312 | 1 | GUB2_HORVU | P12257 hordeum vul |
| 728 | 6 | 0.6 | 291 | 1 | KPRS_METKA | Q8tut6 methanopyru | 801 | 6 | 0.6 | 312 | 1 | OLFI_HUMAN | O43749 homo sapien |
| 729 | 6 | 0.6 | 292 | 1 | ARGB_METMA | Q8pxj8 methanosarc | 802 | 6 | 0.6 | 312 | 1 | OXYR_STRVD | Q9x5p2 streptomyc |
| 730 | 6 | 0.6 | 293 | 1 | PRMA_ECOLI | P28637 escherichia | 803 | 6 | 0.6 | 312 | 1 | RPOA_LACLA | Q9cdy3 lactococcus |
| 731 | 6 | 0.6 | 294 | 1 | LEP_MYCTU | Q10789 mycobacteri | 804 | 6 | 0.6 | 312 | 1 | T2R6_MOUSE | P59530 mus musculu |
| 732 | 6 | 0.6 | 294 | 1 | YHDJ_ECOLI | P28638 escherichia | 805 | 6 | 0.6 | 313 | 1 | OYEG_HUMAN | Q96rd3 homo sapien |
| 733 | 6 | 0.6 | 295 | 1 | ISPE_PASMU | P57833 pasteurella | 806 | 6 | 0.6 | 313 | 1 | YCGC_ECOLI | P55140 escherichia |
| 734 | 6 | 0.6 | 296 | 1 | MY88_HUMAN | Q98336 homo sapien | 807 | 6 | 0.6 | 313 | 1 | YCO6_YEAST | O12094 saccharomyc |
| 735 | 6 | 0.6 | 296 | 1 | YC32_METUA | Q58629 methanococ | 808 | 6 | 0.6 | 314 | 1 | BLA2_STRCI | P14560 streptomyc |
| 736 | 6 | 0.6 | 297 | 1 | BPHC_BURCE | P47228 burkholderi | 809 | 6 | 0.6 | 314 | 1 | FMT_EUCAI | P57584 buchiera ap |
| 737 | 6 | 0.6 | 297 | 1 | HIS1_YEAST | P00498 saccharomyc | 810 | 6 | 0.6 | 314 | 1 | HCDH_HUMAN | Q16836 homo sapien |
| 738 | 6 | 0.6 | 297 | 1 | ISPE_LEPIN | Q8ezm8 leptospira | 811 | 6 | 0.6 | 314 | 1 | SERB_HAEIN | P44997 haemophilus |
| 739 | 6 | 0.6 | 297 | 1 | PRCM_ERWCH | P42194 erwilia chr | 812 | 6 | 0.6 | 315 | 1 | YZ09_AQUAE | O66403 aquifex aeo |
| 740 | 6 | 0.6 | 297 | 1 | Y060_MYCCE | P47306 mycoplasma | 813 | 6 | 0.6 | 316 | 1 | ARGI_XENLA | P30759 xenopus lae |
| 741 | 6 | 0.6 | 298 | 1 | CGIC_RAT | P39947 rattus norv | 814 | 6 | 0.6 | 316 | 1 | KHSE_MYCTU | Q10603 mycobacteri |
| 742 | 6 | 0.6 | 298 | 1 | YHAJ_ECOLI | P42623 escherichia | 815 | 6 | 0.6 | 316 | 1 | MRAW_VIBCH | Q9kpf9 vibrio chol |
| 743 | 6 | 0.6 | 298 | 1 | YIHU_ECOLI | P32142 escherichia | 816 | 6 | 0.6 | 316 | 1 | MRAW_VIBPR | Q9aj99 vibrio prot |
| 744 | 6 | 0.6 | 299 | 1 | ISPE_AGR75 | Q8uhp8 agrobacteri | 817 | 6 | 0.6 | 316 | 1 | MRAW_VIBU | Q8dek2 erwinia chr |
| 745 | 6 | 0.6 | 299 | 1 | ISPE_STRCO | Q9k3r6 streptomyc | 818 | 6 | 0.6 | 316 | 1 | PECT_ERWCH | P52662 erwinia chr |
| 746 | 6 | 0.6 | 299 | 1 | SRK_BACSU | O05510 bacillus su | 819 | 6 | 0.6 | 316 | 1 | TNPI_HUMAN | Q13829 homo sapien |
| 747 | 6 | 0.6 | 299 | 1 | YF02_MYCTU | P71783 mycobacteri | 820 | 6 | 0.6 | 316 | 1 | YNF8_YEAST | P53947 saccharomyc |
| 748 | 6 | 0.6 | 300 | 1 | Y921_METUA | Q58331 methanococ | 821 | 6 | 0.6 | 317 | 1 | FMT_BACSU | P94463 bacillus su |
| 749 | 6 | 0.6 | 301 | 1 | FPQ2_RHIME | Q59752 rhizobium m | 822 | 6 | 0.6 | 317 | 1 | FMT_FUSNN | Q8rdm3 fusobacteri |
| 750 | 6 | 0.6 | 301 | 1 | HWGL_HAEIN | P45099 haemophilus | 823 | 6 | 0.6 | 317 | 1 | PSRC_WOLSU | P31077 wolinnella s |
| 751 | 6 | 0.6 | 301 | 1 | HWGL_PSENV | P13703 pseudomonas | 824 | 6 | 0.6 | 317 | 1 | RL0A_ARATH | O04204 arabidopsis |
| 752 | 6 | 0.6 | 301 | 1 | PRD2_THETN | Q8r9r7 thermoanaer | 825 | 6 | 0.6 | 318 | 1 | CLH2_ARATH | Q9m717 arabidopsis |
| 753 | 6 | 0.6 | 302 | 1 | BPHC_PSEPS | P08695 pseudomonas | 826 | 6 | 0.6 | 318 | 1 | OYDI_HUMAN | Q8m9x2 chaetosphe |
| 754 | 6 | 0.6 | 302 | 1 | PR12_DROME | P12982 drosophila | 827 | 6 | 0.6 | 318 | 1 | OYDI_HUMAN | Q9h346 homo sapien |
| 755 | 6 | 0.6 | 302 | 1 | PR13_DROME | Q05547 drosophila | 828 | 6 | 0.6 | 318 | 1 | T2R7_HUMAN | Q9nyw3 homo sapien |
| 756 | 6 | 0.6 | 302 | 1 | RBCR_CHRVI | P25544 chromatium | 829 | 6 | 0.6 | 318 | 1 | YCCG_BACSU | P70954 bacillus su |
| 757 | 6 | 0.6 | 303 | 1 | CGIC_HUMAN | P24863 homo sapien | 830 | 6 | 0.6 | 319 | 1 | HEMZ_STRMU | Q8cmw4 streptococ |
| 758 | 6 | 0.6 | 303 | 1 | MTAA_SYNY3 | P74040 synechocyst | 831 | 6 | 0.6 | 319 | 1 | PTET_HUMAN | Q14734 homo sapien |
| 759 | 6 | 0.6 | 303 | 1 | T2RD_HUMAN | Q9nyv9 homo sapien | 832 | 6 | 0.6 | 319 | 1 | RLA0_ORYSA | P41095 oryza sativ |
| 760 | 6 | 0.6 | 304 | 1 | CGIC_MOUSE | Q62447 mus musculu | 833 | 6 | 0.6 | 320 | 1 | ASPG_SPOFR | Q02467 spodoptera |
| 761 | 6 | 0.6 | 304 | 1 | SUAC_RAT | P50237 rattus norv | 834 | 6 | 0.6 | 320 | 1 | PTET_MOUSE | P58137 mus musculu |
| 762 | 6 | 0.6 | 305 | 1 | ARCC_THEVO | Q97c45 thermoplas | 835 | 6 | 0.6 | 321 | 1 | EAEB_ECO27 | Q05129 escherichia |
| 763 | 6 | 0.6 | 305 | 1 | ERA_PSEAE | Q9xcx8 pseudomonas | 836 | 6 | 0.6 | 321 | 1 | 1F2B_SCHPO | P56329 schizosacch |

| | | | | | | | | | | | | | | | |
|-----|---|-----|-----|---|-------------|--------|--------------|-----|---|-----|-----|---|------------|--------|---------------|
| 837 | 6 | 0.6 | 321 | 1 | RBSC_ECOLI | P04984 | escherichia | 910 | 6 | 0.6 | 331 | 1 | MACS_BOVIN | P12624 | bos taurus |
| 838 | 6 | 0.6 | 321 | 1 | SPX3_HUMAN | Q9bwm7 | homo sapien | 911 | 6 | 0.6 | 331 | 1 | PSO2_ARATH | Q9841 | arabidopsis |
| 839 | 6 | 0.6 | 321 | 1 | SPX3_MOUSE | Q91v61 | mus musculus | 912 | 6 | 0.6 | 331 | 1 | RL3_ARCFU | Q28354 | archaeoglob |
| 840 | 6 | 0.6 | 321 | 1 | SPX3_RAT | Q91jy2 | rattus norv | 913 | 6 | 0.6 | 331 | 1 | RPOS_YEREN | P47765 | yersinia en |
| 841 | 6 | 0.6 | 322 | 1 | ARGI_HUMAN | P05089 | homo sapien | 914 | 6 | 0.6 | 331 | 1 | YB30_ARATH | O82114 | arabidopsis |
| 842 | 6 | 0.6 | 322 | 1 | GRP2_MOUSE | O89100 | m grb2-rela | 915 | 6 | 0.6 | 332 | 1 | HRDD_STRCO | P18249 | streptomyce |
| 843 | 6 | 0.6 | 322 | 1 | MIAA_WIGBR | O8d318 | wiggleswort | 916 | 6 | 0.6 | 332 | 1 | HRDD_STRVF | Q60012 | streptomyce |
| 844 | 6 | 0.6 | 322 | 1 | NADA_CHLTE | Q8kew2 | chlorobium | 917 | 6 | 0.6 | 332 | 1 | HS30_YEAST | P25619 | saccharomyc |
| 845 | 6 | 0.6 | 322 | 1 | FOR_SYNYT3 | Q59987 | synecocyst | 918 | 6 | 0.6 | 332 | 1 | PSBO_TOBAC | O40459 | nicotiana t |
| 846 | 6 | 0.6 | 322 | 1 | RLAO_LUPLU | P50345 | lupinus lut | 919 | 6 | 0.6 | 332 | 1 | PSOI_ARATH | P23321 | arabidopsis |
| 847 | 6 | 0.6 | 322 | 1 | SFX1_HUMAN | Q9h9b4 | homo sapien | 920 | 6 | 0.6 | 333 | 1 | PSBO_SOLTU | P26320 | solanum tub |
| 848 | 6 | 0.6 | 322 | 1 | SFX1_MOUSE | Q99j11 | mus musculus | 921 | 6 | 0.6 | 333 | 1 | VINT_BPMF2 | Q83361 | mycobacteri |
| 849 | 6 | 0.6 | 322 | 1 | SFX1_RAT | Q63955 | rattus norv | 922 | 6 | 0.6 | 333 | 1 | VINT_BPMF2 | P25426 | mycobacteri |
| 850 | 6 | 0.6 | 322 | 1 | YQJA_BACSU | P54538 | bacillus su | 923 | 6 | 0.6 | 334 | 1 | CATL_MOUSE | P06797 | mus musculus |
| 851 | 6 | 0.6 | 323 | 1 | ARGI_MOUSE | Q61176 | mus musculus | 924 | 6 | 0.6 | 334 | 1 | CATL_MOUSE | P07154 | rattus norv |
| 852 | 6 | 0.6 | 323 | 1 | ARGI_RAT | P07824 | rattus norv | 925 | 6 | 0.6 | 334 | 1 | UNG_HSV11 | P10186 | herpes simp |
| 853 | 6 | 0.6 | 323 | 1 | HEMZ_HAEIN | P43868 | haemophilus | 926 | 6 | 0.6 | 334 | 1 | Y472_RICPR | Q92d72 | ricketsia |
| 854 | 6 | 0.6 | 323 | 1 | VP9_RGDV | P23628 | rice gall d | 927 | 6 | 0.6 | 335 | 1 | ILC1_SULSO | Q9uwx9 | sulfolobus |
| 855 | 6 | 0.6 | 324 | 1 | CAHE_MOUSE | Q99n23 | mus musculus | 928 | 6 | 0.6 | 335 | 1 | OTCC_PSEAE | P08308 | pseudomonas |
| 856 | 6 | 0.6 | 324 | 1 | ODPB_BACSU | P21882 | bacillus su | 929 | 6 | 0.6 | 335 | 1 | RPOS_VIBCH | O51804 | vibrio chol |
| 857 | 6 | 0.6 | 324 | 1 | YL70_ARCFU | O28112 | archaeoglob | 930 | 6 | 0.6 | 335 | 1 | STSI_ARATH | P94111 | arabidopsis |
| 858 | 6 | 0.6 | 325 | 1 | ACC2_DORSP | Q39705 | doritaenops | 931 | 6 | 0.6 | 336 | 1 | ASNA_CLOPE | Q8x1y5 | clostridium |
| 859 | 6 | 0.6 | 325 | 1 | LACD_STRMU | P26425 | streptococc | 932 | 6 | 0.6 | 336 | 1 | ID12_STRPN | Q97488 | streptococc |
| 860 | 6 | 0.6 | 325 | 1 | PSBO_WHEAT | P27665 | trititum ae | 933 | 6 | 0.6 | 336 | 1 | LEU3_SULTO | P50455 | sulfolobus |
| 861 | 6 | 0.6 | 325 | 1 | RPOA_MYCGA | Q9rdv6 | mycoplasma | 934 | 6 | 0.6 | 336 | 1 | MRAY_STRP3 | Q8k6c7 | streptococc |
| 862 | 6 | 0.6 | 325 | 1 | Y132_CHLPN | Q92949 | chlamydia p | 935 | 6 | 0.6 | 336 | 1 | MRAY_STRP3 | Q8n3y2 | streptococc |
| 863 | 6 | 0.6 | 326 | 1 | HEMZ_PASMU | P57874 | pasteurella | 936 | 6 | 0.6 | 336 | 1 | MRAY_STRP3 | Q8n3y2 | streptococc |
| 864 | 6 | 0.6 | 326 | 1 | LACD_LACTA | P26593 | lactococcus | 937 | 6 | 0.6 | 336 | 1 | MRAY_STRP3 | Q99y42 | streptococc |
| 865 | 6 | 0.6 | 326 | 1 | LACD_STAAM | P11100 | staphylococ | 938 | 6 | 0.6 | 337 | 1 | RPOA_ORISA | P12090 | oryza sativ |
| 866 | 6 | 0.6 | 326 | 1 | NULM_CHOCR | P48898 | chondrus cr | 939 | 6 | 0.6 | 337 | 1 | TRPD_BRAJA | P94326 | bradyrhizob |
| 867 | 6 | 0.6 | 326 | 1 | VS09_ROTPEL | P25176 | equine rota | 940 | 6 | 0.6 | 338 | 1 | YD43_SCHPO | P87052 | schizosacch |
| 868 | 6 | 0.6 | 326 | 1 | VS09_ROTPEM | P32549 | porcine rot | 941 | 6 | 0.6 | 339 | 1 | LPXD_YERPE | P58611 | yersinia pe |
| 869 | 6 | 0.6 | 327 | 1 | CCG4_HUMAN | Q9ubn1 | homo sapien | 942 | 6 | 0.6 | 339 | 1 | PAX9_CHICK | P55166 | gallus gall |
| 870 | 6 | 0.6 | 327 | 1 | CCG4_MOUSE | Q91jv4 | mus musculus | 943 | 6 | 0.6 | 339 | 1 | RE13_YEAST | P12686 | saccharomyc |
| 871 | 6 | 0.6 | 327 | 1 | CCG4_RAT | Q8vhw9 | rattus norv | 944 | 6 | 0.6 | 339 | 1 | YE60_XYLFA | Q9pdb9 | xylorella fas |
| 872 | 6 | 0.6 | 327 | 1 | FKB6_HUMAN | O75344 | homo sapien | 945 | 6 | 0.6 | 340 | 1 | THPS_SULAC | P17118 | sulfolobus |
| 873 | 6 | 0.6 | 327 | 1 | ODPB_ACHLA | P35488 | acholeplasm | 946 | 6 | 0.6 | 340 | 1 | YM21_ARCFU | O28062 | archaeoglob |
| 874 | 6 | 0.6 | 327 | 1 | PP1L_DROME | P48461 | drosophila | 947 | 6 | 0.6 | 341 | 1 | MDHM_CAEEL | O02640 | caenorhabdi |
| 875 | 6 | 0.6 | 327 | 1 | PSTC_RHLIO | Q98f13 | rhizobium l | 948 | 6 | 0.6 | 341 | 1 | MRW_RHIME | Q22414 | rhizobium m |
| 876 | 6 | 0.6 | 327 | 1 | TAL_CHLMU | Q9pk80 | chlamydia m | 949 | 6 | 0.6 | 341 | 1 | OMPU_VIBCH | P97085 | vibrio chol |
| 877 | 6 | 0.6 | 327 | 1 | TAL_CHLTR | O84315 | chlamydia t | 950 | 6 | 0.6 | 341 | 1 | POLG_HPAV1 | P13672 | hepatitis a |
| 878 | 6 | 0.6 | 328 | 1 | HAM1_STRP3 | Q8k817 | streptococc | 951 | 6 | 0.6 | 342 | 1 | CITC_KLEPN | P45410 | klebsiella |
| 879 | 6 | 0.6 | 328 | 1 | HAM1_STRP8 | Q8p242 | streptococc | 952 | 6 | 0.6 | 342 | 1 | LYCA_BPCP7 | P19385 | bacterioph |
| 880 | 6 | 0.6 | 328 | 1 | HAM1_STRPY | Q9a1b6 | streptococc | 953 | 6 | 0.6 | 342 | 1 | NMT1_ASPPA | P42882 | aspergillus |
| 881 | 6 | 0.6 | 328 | 1 | I12B_HUMAN | P29460 | homo sapien | 954 | 6 | 0.6 | 342 | 1 | TRM1_THEVO | Q97ar2 | thermoplasm |
| 882 | 6 | 0.6 | 328 | 1 | I12B_MACMU | P48095 | macaca mula | 955 | 6 | 0.6 | 343 | 1 | HRCA_BACSU | P25499 | bacillus su |
| 883 | 6 | 0.6 | 328 | 1 | NAS6_HORVU | P75476 | mycoplasma | 956 | 6 | 0.6 | 343 | 1 | SRT1_YEAST | Q03175 | saccharomyc |
| 884 | 6 | 0.6 | 328 | 1 | NODZ_AZOCA | Q9zgv3 | hordeum vul | 957 | 6 | 0.6 | 343 | 1 | VIBY_AGRTU | P05360 | agrobacteri |
| 885 | 6 | 0.6 | 328 | 1 | SYFA_EUCBP | Q4366 | azorhizobiu | 958 | 6 | 0.6 | 344 | 1 | CGG2_MOUSE | O08918 | mus musculus |
| 886 | 6 | 0.6 | 328 | 1 | YU77_YEAST | P59504 | buchnera ap | 959 | 6 | 0.6 | 344 | 1 | COA2_POVJC | P03095 | polyomaviru |
| 887 | 6 | 0.6 | 328 | 1 | YRHB_ECOLI | P47145 | saccharomyc | 960 | 6 | 0.6 | 344 | 1 | FLIM_TREPA | P74927 | treponema p |
| 888 | 6 | 0.6 | 328 | 1 | YRHB_ECOLI | P45395 | escherichia | 961 | 6 | 0.6 | 344 | 1 | NU2M_RANCA | P16673 | rana catesb |
| 889 | 6 | 0.6 | 329 | 1 | CAHX_FLAPR | P46281 | flaveria pr | 962 | 6 | 0.6 | 344 | 1 | SYW_CHLPN | Q9z7a4 | chlamydia p |
| 890 | 6 | 0.6 | 329 | 1 | ID12_STRPY | Q9a095 | streptococc | 963 | 6 | 0.6 | 345 | 1 | D2D2_XENLA | P34973 | xenopus lae |
| 891 | 6 | 0.6 | 329 | 1 | NAS4_HORVU | Q9zgv6 | hordeum vul | 964 | 6 | 0.6 | 346 | 1 | CDK7_HUMAN | P05613 | homo sapien |
| 892 | 6 | 0.6 | 329 | 1 | NAS7_HORVU | Q9zwh8 | hordeum vul | 965 | 6 | 0.6 | 346 | 1 | EPRI_AERHY | O05485 | aeromonas h |
| 893 | 6 | 0.6 | 329 | 1 | NAS8_HORVU | Q9xfb6 | hordeum vul | 966 | 6 | 0.6 | 346 | 1 | GALM_ECOLI | O05485 | aeromonas h |
| 894 | 6 | 0.6 | 329 | 1 | PDXA_NEIMA | Q9jx42 | neisseria m | 967 | 6 | 0.6 | 346 | 1 | PSTY_ECOLI | P40681 | escherichia |
| 895 | 6 | 0.6 | 329 | 1 | PDXA_NEIMB | Q9klf9 | neisseria m | 968 | 6 | 0.6 | 346 | 1 | SYW_CHLMU | P06128 | escherichia |
| 896 | 6 | 0.6 | 329 | 1 | PSBO_FRIAG | O49079 | fritillaria | 969 | 6 | 0.6 | 346 | 1 | SYW_CHLTR | Q9pjf5 | chlamydia m |
| 897 | 6 | 0.6 | 329 | 1 | PSBO_LYCES | P23322 | lycopersico | 970 | 6 | 0.6 | 346 | 1 | TAS_ECOLI | O46933 | escherichia |
| 898 | 6 | 0.6 | 329 | 1 | PSBO_FEA | P14226 | pisum sativ | 971 | 6 | 0.6 | 346 | 1 | YE77_METJA | Q58872 | methanococc |
| 899 | 6 | 0.6 | 329 | 1 | SRPG_SYN7P | Q59966 | synecococc | 972 | 6 | 0.6 | 347 | 1 | ITP2_HUMAN | Q58872 | methanococc |
| 900 | 6 | 0.6 | 329 | 1 | Y461_CHLNR | O84467 | chlamydia t | 973 | 6 | 0.6 | 347 | 1 | NADA_ECO57 | Q8x963 | escherichia |
| 901 | 6 | 0.6 | 329 | 1 | YCLC_PSEUP | Q47100 | pseudomonas | 974 | 6 | 0.6 | 347 | 1 | NADA_ECOLI | P11458 | escherichia |
| 902 | 6 | 0.6 | 330 | 1 | CAH1_FLALI | P46512 | flaveria li | 975 | 6 | 0.6 | 347 | 1 | NADA_SALTI | Q82858 | salmonella |
| 903 | 6 | 0.6 | 330 | 1 | CAHX_FLABR | P46511 | flaveria br | 976 | 6 | 0.6 | 347 | 1 | NADA_SALTY | P24519 | salmonella |
| 904 | 6 | 0.6 | 330 | 1 | GRP2_HUMAN | O75791 | h grb2-rela | 977 | 6 | 0.6 | 347 | 1 | NADA_SHIFL | Q93cp9 | shigella fl |
| 905 | 6 | 0.6 | 330 | 1 | PEXD_CAEEL | Q19951 | caenorhabdi | 978 | 6 | 0.6 | 348 | 1 | STRI_STRGR | P09400 | streptomyce |
| 906 | 6 | 0.6 | 330 | 1 | RPOS_ECOLI | P13445 | escherichia | 979 | 6 | 0.6 | 349 | 1 | ESR1_ANOXA | Q9ynt3 | anolis caro |
| 907 | 6 | 0.6 | 330 | 1 | RPOS_SALTY | P37400 | salmonella | 980 | 6 | 0.6 | 349 | 1 | LPXD_ANASP | O8y810 | anaeaba sp |
| 908 | 6 | 0.6 | 330 | 1 | RPOS_SHIFL | P35540 | ehigella fl | 981 | 6 | 0.6 | 350 | 1 | MRP_DEIRA | Q9rvm9 | deinococcus |
| 909 | 6 | 0.6 | 331 | 1 | CAHX_FLABI | P46510 | flaveria bi | 982 | 6 | 0.6 | 350 | 1 | VOD2_DROME | Q9vcq3 | drosophila |

983 P76349 escherichia
 984 P14997 polyomaviru
 985 P03094 polyomaviru
 986 Q32861 mycoplasma
 987 Q98w5 xanthomonas
 988 Q97w02 sulfobolus
 989 Q57511 borrelia bu
 990 P72183 propionibac
 991 O68547 rhizobium l
 992 P47793 brachydanio
 993 P26189 salmonella
 994 O02754 bos taurus
 995 Q8kcq3 chlorobium
 996 Q829v8 yersinia pe
 997 O34472 bacillus su
 998 P43931 haemophilus
 999 Q60739 mus musculu
 1000 Q8y7b6 listeria mo

ALIGNMENTS

RESULT 1
 ID_HLY1_ECOLI STANDARD; PRT; 1023 AA.
 AC P09983;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hemolysin, chromosomal.
 GN HLYA.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J96 / Serotype O4;
 RX MEDLINE=85234404; PubMed=3891743;
 RA Felmelee T., Pellett S., Welch R.A.;
 RT "Nucleotide sequence of an Escherichia coli chromosomal hemolysin.";
 RL J. Bacteriol. 163:94-105(1985).
 RN [2]
 RP SEQUENCE OF 1-44 FROM N.A.
 RC STRAIN=2001;
 RX MEDLINE=85258115; PubMed=3894051;
 RA Nicaud J.-M., Mackman N., Gray L., Holland I.B.;
 RT "Characterisation of HlyC and mechanism of activation and secretion
 of haemolysin from E. coli 2001.";
 RL FEBS Lett. 187:339-344(1985).
 CC -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
 cell membranes and cause cell rupture by mechanisms not clearly
 defined.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- DOMAIN: The gly-rich region is probably involved in binding
 calcium, which is required for target cell-binding or cytolytic
 activity.
 CC -!- DOMAIN: The three transmembrane domains are believed to be
 involved in pore formation by the cytotoxin.
 CC -!- PTM: Palmitoylated by hlyC. The toxin only becomes active when
 modified.
 CC -!- MISCELLANEOUS: The hemolysin of E.coli is produced predominantly
 by strains causing extraintestinal infections, such as those of
 the urinary tract.
 CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

 CC EMBL; M10133; AAA23975.1; --
 CC EMBL; X02768; CAA26546.1; --
 DR PIR; A24433; LEECA.
 DR InterPro; IPR001343; Hemolysin_Ca_bind.
 DR InterPro; IPR003995; RTX.
 DR Pfam; PF00353; hemolysinCbind; 6.
 DR Pfam; PF02382; RTX_1.
 DR PRINTS; PRO0313; CABDNGRPT.
 DR PRINTS; PRO1488; RTXTOXINA.
 DR PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
 KW Hemolysin; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
 KW Lipoprotein; Palmitate.
 FT TRANSMEM 237 259 POTENTIAL.
 FT TRANSMEM 267 326 POTENTIAL.
 FT TRANSMEM 364 410 POTENTIAL.
 FT DOMAIN 723 869 16 X REPEATS, GLY-RICH.
 FT REPEAT 723 728 1.
 FT REPEAT 732 737 2.
 FT REPEAT 741 746 3.
 FT REPEAT 750 755 4.
 FT REPEAT 759 764 5.
 FT REPEAT 768 773 6.
 FT REPEAT 777 782 7.
 FT REPEAT 786 791 8.
 FT REPEAT 795 800 9.
 FT REPEAT 806 812 10.
 FT REPEAT 821 821 11.
 FT REPEAT 825 830 12.
 FT REPEAT 834 839 13.
 FT REPEAT 843 848 14.
 FT REPEAT 855 860 15.
 FT REPEAT 864 869 16.
 FT LIPID 563 563 PALMITATE (BY SIMILARITY).
 FT LIPID 689 689 PALMITATE (BY SIMILARITY).
 FT VARIANT 6 A -> T (IN STRAIN 2001).
 SQ SEQUENCE 1023 AA; 109867 MW; 196D5C0A9A28B54D CRC64;

Query Match 1.5%; Score 14; DB 1; Length 1023;
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 379 ISGILEASKOAMPE 392
 Db 405 ISGILEASKOAMPE 418
 |||||

RESULT 2
 ID_HLYA_ECOLI STANDARD; PRT; 1024 AA.
 AC P08715;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hemolysin, plasmid.
 GN HLYA.
 OS Escherichia coli.
 OG Plasmid Inci2 pHLy152.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hess J., Wells W., Vogel M., Goebel W.;
 RT "Nucleotide sequence of a plasmid-encoded hemolysin determinant and
 its comparison with a corresponding chromosomal hemolysin sequence.";
 RL FEMS Microbiol. Lett. 34:1-11(1986).
 RN [2]
 RP PALMITOYLATION OF LYS-564 AND LYS-690.
 RX MEDLINE=9509325; PubMed=7801126;
 RA Stanley P., Packman L.C., Koronakis V., Hughes C.;
 RT "Patty acylation of two internal lysine residues required for the

RT toxic activity of *Escherichia coli* hemolysin."; Science 266:1992-1996(1994).
 RN [3]
 RP PALMITOYLATION OF LYS-564 AND LYS-690.
 RA MEDLINE=96040790; PubMed=8808931;
 RX Ludwig A., Garcia F., Jarchau T., Benz R., Hoppe J.,
 RA Goebel W.;
 RT "Analysis of the in vivo activation of hemolysin (HlyA) from
 RT *Escherichia coli*.";
 RL J. Bacteriol. 178:5422-5430(1996).
 CC -1- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
 CC cell membranes and cause cell rupture by mechanisms not clearly
 CC defined.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DOMAIN: The Gly-rich region is probably involved in binding
 CC calcium, which is required for target cell-binding or cytolytic
 CC activity.
 CC -1- DOMAIN: The three transmembrane domains are believed to be
 CC involved in pore formation by the cytotoxin.
 CC -1- PTM: Palmitoylated by hlyC. The toxin only becomes active when
 CC modified.
 CC -1- MISCELLANEOUS: The hemolysin of *E. coli* is produced predominantly
 CC by strains causing extraintestinal infections, such as those of
 CC the urinary tract.
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M14107; AAA98233.1; -;
 DR InterPro; IPR001343; Hemolysin_Ca_bind.
 DR InterPro; IPR003995; RtxA.
 DR Pfam; PF00353; hemolysinCabin; 6.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PR00313; CABNDNGRPT.
 DR PRINTS; PR01488; RTXTOXINA.
 DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 4.
 DR Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
 KW Lipoprotein; Palmitate; Plasmid;
 FT TRANSMEM 238 260 POTENTIAL.
 FT TRANSMEM 268 327 POTENTIAL.
 FT TRANSMEM 365 411 POTENTIAL.
 FT DOMAIN 724 870 16 X REPEATS, GLY-RICH.
 FT REPEAT 724 729 1.
 FT REPEAT 733 738 2.
 FT REPEAT 742 747 3.
 FT REPEAT 751 756 4.
 FT REPEAT 760 765 5.
 FT REPEAT 769 774 6.
 FT REPEAT 778 783 7.
 FT REPEAT 787 792 8.
 FT REPEAT 796 801 9.
 FT REPEAT 807 812 10.
 FT REPEAT 817 822 11.
 FT REPEAT 826 831 12.
 FT REPEAT 835 840 13.
 FT REPEAT 844 849 14.
 FT REPEAT 856 861 15.
 FT REPEAT 865 870 16.
 FT LIPID 564 564 PALMITATE.
 FT LIPID 690 690 PALMITATE.
 SQ SEQUENCE 1024 AA; 110201 NW; 83944917F76C945B CRC64;
 Query Match 1.5%; Score 14; DB 1; Length 1024;
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 379 ISGILEASKQAMPE 392

Db 406 ISGILEASKQAMFE 419
 |||||
 RESULT 3
 RT31_ACTPL STANDARD; PRT; 1049 AA.
 ID RT31_ACTPL
 AC P55130;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE RTX-III toxin determinant A from serotype 2 (APX-IIIA) (Cytolysin
 DE IIIA) (CLY-IIIA).
 GN APXIIIA OR CLYIIIA OR RTXIA OR PTXA.
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Actinobacillus.
 OX NCBI_TaxID=715;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Serotype 2;
 RX MEDLINE=93263992; PubMed=8494611;
 RA Chang Y.-F., Shi J., Ma D.-P., Shin S.J., Lein D.H.;
 RT "Molecular analysis of the Actinobacillus pleuropneumoniae RTX
 RT toxin-III gene cluster.";
 RL DNA Cell Biol. 12:351-362(1993).
 RN [2]
 RP SEQUENCE OF 828-1049 FROM N.A.
 RC STRAIN=1536 / Serotype 2;
 RX MEDLINE=95012630; PubMed=7927703;
 RA Jansen R., Braire J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J.,
 RA Smits M.A.;
 RT "Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (ApX)
 RT operons: Characterization of the ApXIII operons.";
 RL Infect. Immun. 62:4411-4418(1994).
 CC -1- FUNCTION: Does not have hemolytic activity but shows a strong
 CC cytotoxicity towards alveolar macrophages and neutrophils.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DOMAIN: The Gly-rich region is probably involved in binding
 CC calcium, which is required for target cell-binding or cytolytic
 CC activity (By similarity).
 CC -1- DOMAIN: The three transmembrane domains are believed to be
 CC involved in pore formation by the cytotoxin (By similarity).
 CC -1- PTM: Palmitoylated by apXIIIC. The toxin only becomes active when
 CC modified (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L12145; AAA21924.1; -;
 DR EMBL; X80056; CAB37652.1; ALT_SEQ.
 DR PIR; S51784; S51784.
 DR InterPro; IPR001343; Hemolysin_Ca_bind.
 DR InterPro; IPR003995; RtxA.
 DR Pfam; PF00353; hemolysinCabin; 6.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PR00313; CABNDNGRPT.
 DR PRINTS; PR01488; RTXTOXINA.
 DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 3.
 KW Toxin; Cytolysis; Repeat; Calcium; Transmembrane; Lipoprotein;
 KW Palmitate.
 FT TRANSMEM 154 170 POTENTIAL.
 FT TRANSMEM 315 331 POTENTIAL.
 FT TRANSMEM 397 413 POTENTIAL.
 FT DOMAIN 753 858 11 X REPEATS, GLY-RICH.
 FT REPEAT 753 758 1.
 FT REPEAT 762 767 2.

```

FT REPEAT      771 776      3.
FT REPEAT      780 785      4.
FT REPEAT      789 794      5.
FT REPEAT      798 803      6.
FT REPEAT      807 812      7.
FT REPEAT      826 831      8.
FT REPEAT      835 840      9.
FT REPEAT      844 849     10.
FT REPEAT      853 858     11.
SQ SEQUENCE 1049 AA; 112491 MW; F99846BFD4E5CE72 CRC64;

Query Match      1.5%; Score 14; DB 1; Length 1049;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 KVAAGFELSNOVIG 254
Db 275 KVAAGFELSNOVIG 288

RESULT 4
RT32 ACTPL STANDARD; PRT; 1052 AA.
AC P55131;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE RTX-III toxin determinant A from serotype 8 (APX-IIIa) (Cytolysin
  IIa) (CLY-IIIa).
GN APXIIIA OR CLYIIIA OR RTXA OR PTXA.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=715;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=405 / Serotype 8;
RX MEDLINE=95012630; PubMed=7927703;
RA Jansen R., Briaire J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J.,
  Smits M.A.;
RT "Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (ApX)
  operons: characterization of the ApXIII operons."
RL Infect. Immun. 62:4411-4418(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype 8;
RX MEDLINE=93162836; PubMed=8432615;
RA Jansen R., Briaire J., Kamp E.M., Gielkens A.L.J., Smits M.A.;
RT "Cloning and characterization of the Actinobacillus
  pleuropneumoniae-RTX-toxin III (ApXIII) gene."
RL Infect. Immun. 61:947-954(1993).
CC -!- FUNCTION: Does not have hemolytic activity but shows a strong
  cytotoxicity towards alveolar macrophages and neutrophils.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: The gly-rich region is probably involved in binding
  calcium, which is required for target cell-binding or cytolytic
  activity (By similarity).
CC -!- DOMAIN: The three transmembrane domains are believed to be
  involved in pore formation by the cytotoxin (By similarity).
CC -!- PTM: Palmitoylated by apXIIIC. The toxin only becomes active when
  modified (By similarity).
CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X80055; CAA56358.1; -.
CC EMBL; X68815; CAA48711.1; -.

```

```

DR PIR; B49219; B49219.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCbind; 6.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTXINA.
DR PROSITE; PS00330; HEMOLYSIN CALCIUM; 3.
KW Toxin; Cytolysis; Repeat; Calcium; Transmembrane; Lipoprotein;
KW Palmitate.
FT TRANSMEM 248 265 POTENTIAL.
FT TRANSMEM 275 334 POTENTIAL.
FT TRANSMEM 372 418 POTENTIAL.
FT DOMAIN 754 859 7 X REPEATS, GLY-RICH.
FT REPEAT 754 759 1.
FT REPEAT 763 768 2.
FT REPEAT 772 777 3.
FT REPEAT 781 786 4.
FT REPEAT 790 795 5.
FT REPEAT 808 813 6.
FT REPEAT 827 832 7.
FT REPEAT 836 841 8.
FT REPEAT 845 850 9.
FT REPEAT 854 859 10.
FT REPEAT 854 859 11.
SQ SEQUENCE 1052 AA; 112809 MW; F83AFE25A6FD8758 CRC64;

Query Match      1.5%; Score 14; DB 1; Length 1052;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 KVAAGFELSNOVIG 254
Db 275 KVAAGFELSNOVIG 288

RESULT 5
HLVA ACTSU STANDARD; PRT; 956 AA.
AC Q00951;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hemolysin (Cytolysin II) (CLY-IIa) (HLX-IIa) (CYTC) (APPA).
GN APPA OR CLYIIA OR HLXIIA OR CYTC.
OS Actinobacillus suis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=716;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3714;
RX MEDLINE=92267623; PubMed=1587585;
RA Burrows L.L., Lo R.Y.;
RT "Molecular characterization of an RTX toxin determinant from
  Actinobacillus suis."
RL Infect. Immun. 60:2166-2173(1992).
CC -!- FUNCTION: One of the virulence factors of A.suis might be a
  secreted cytotoxin, possibly the extracellular hemolysin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: The gly-rich region is probably involved in binding
  calcium, which is required for target cell-binding or cytolytic
  activity.
CC -!- DOMAIN: The three transmembrane domains are believed to be
  involved in pore formation by the cytotoxin (By similarity).
CC -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
  modified (By similarity).
CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X80055; CAA56358.1; -.
CC EMBL; X68815; CAA48711.1; -.

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M90440; AAA21918.1; -
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR Pfam; PF00353; hemolysinCabin; 5.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PRO0313; CABNDNGRPT.
 DR PRINTS; PRO1488; RTXTOXINA.
 DR PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
 KW Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
 KW Lipoprotein; Palmitate.
 FT TRANSMEM 238 254 POTENTIAL.
 FT TRANSMEM 302 320 POTENTIAL.
 FT TRANSMEM 383 406 POTENTIAL.
 FT DOMAIN 719 799 9 X REPEATS, GLY-RICH.
 FT REPEAT 719 724 1.
 FT REPEAT 728 733 2.
 FT REPEAT 737 742 3.
 FT REPEAT 746 751 4.
 FT REPEAT 755 760 5.
 FT REPEAT 764 769 6.
 FT REPEAT 773 778 7.
 FT REPEAT 782 787 8.
 FT REPEAT 794 799 9.
 SQ SEQUENCE 956 AA; 102453 MW; 3415FF1D7ADD4365 CRC64;

Query Match 1.2%; Score 11; DB 1; Length 956;
 Best Local Similarity 100.0%; Pred. No. 0.023;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 369 ALLVAGVTGLI 379
 |||||
 Db 392 ALLVAGVTGLI 402

RESULT 6

RT2A ACTPL RT2A ACTPL STANDARD; PRT; 956 AA.
 AC P15377;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE RTX-II toxin determinant A (APX-IIA) (Hemolysin IIA) (HLY-IIA)
 DE (Cytolysin IIA) (CLY-IIA).
 GN APXIIA OR CLYIIA OR HLYIIA OR APPA OR CVTC.
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Actinobacillus.
 OX NCBI_TaxID=715;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Serotype 5;
 RX MEDLINE=90126233; PubMed=2693022;
 RA Chang Y.-F., Young R., Struck D.K.;
 RT "Cloning and characterization of a hemolysin gene from Actinobacillus
 RT (Haemophilus) pleuropneumoniae.";
 RL DNA 8.635-647(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate CVI 13261 / Serotype 9;
 RX MEDLINE=92040145; PubMed=1937809;
 RA Smits M.A., Briare J., Jansen R., Smith H.E., Kamp E.M.,
 RA Gielkens A.L.;
 RT "Cytolysins of Actinobacillus pleuropneumoniae serotype 9.";
 RL Infect. Immun. 59:4497-4504(1991).
 CC -1- FUNCTION: One of the virulence factors of A. pleuropneumoniae,
 CC which shows a weak hemolytic activity and is moderately cytotoxic
 CC for alveolar macrophages and neutrophils.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DOMAIN: The Gly-rich region is probably involved in binding

CC calcium, which is required for target cell-binding or cytolytic
 CC activity (By similarity).
 CC -1- DOMAIN: The three transmembrane domains are believed to be
 CC involved in pore formation by the cytotoxin (By similarity).
 CC -1- PFM: Palmitoylated by apxIIc. The toxin only becomes active when
 CC modified (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M30602; AAA87232.1; -
 DR EMBL; X61111; CAA43423.1; -
 DR PIR; B33389; B33389.
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR InterPro; IPR003995; RtxA.
 DR Pfam; PF00353; hemolysinCabin; 5.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PRO0313; CABNDNGRPT.
 DR PRINTS; PRO1488; RTXTOXINA.
 DR PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
 KW Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
 KW Lipoprotein; Palmitate.
 FT TRANSMEM 233 256 POTENTIAL.
 FT TRANSMEM 266 323 POTENTIAL.
 FT TRANSMEM 361 406 POTENTIAL.
 FT DOMAIN 719 787 8 X REPEATS, GLY-RICH.
 FT REPEAT 719 724 1.
 FT REPEAT 728 733 2.
 FT REPEAT 737 742 3.
 FT REPEAT 746 751 4.
 FT REPEAT 755 760 5.
 FT REPEAT 764 769 6.
 FT REPEAT 773 778 7.
 FT REPEAT 782 787 8.
 SQ SEQUENCE 956 AA; 102531 MW; BDBCABBADF14A641 CRC64;

Query Match 1.2%; Score 11; DB 1; Length 956;
 Best Local Similarity 100.0%; Pred. No. 0.023;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 369 ALLVAGVTGLI 379
 |||||
 Db 392 ALLVAGVTGLI 402

RESULT 7

RT11 ACTPL RT11 ACTPL STANDARD; PRT; 1023 AA.
 AC P55128;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE RTX-I toxin determinant A from serotypes 1/9 (APX-IA) (Hemolysin IA)
 DE (HLY-IA) (Cytolysin IA) (CLY-IA).
 GN APXIA OR CLYIA OR HLYIA.
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Actinobacillus.
 OX NCBI_TaxID=715;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S 4074 / Serotype 1;
 RX MEDLINE=91348845; PubMed=1879928;
 RA Frey J., Meier R., Gysi D., Nicolet J.;
 RT "Nucleotide sequence of the hemolysin I gene from Actinobacillus
 RT pleuropneumoniae.";
 RT Infect. Immun. 59:3026-3032(1991).
 RL

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S 4074 / Serotype 1;
RX MEDLINE=94237497; PubMed=8181764;
RA Frey J., Haldemann A., Nicolet J., Boffini A., Prentki P.;
RT "Sequence analysis and transcription of the apxI operon (hemolysin I)
RL from Actinobacillus pleuropneumoniae.";
RL Gene 142:97-102(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate CVI 13261 / Serotype 9;
RX MEDLINE=93366425; PubMed=8359891;
RA Jansen R., Briaire J., Kamp E.M., Gielkens A.L.J., Smits M.A.;
RT "Structural analysis of the Actinobacillus pleuropneumoniae-RTX-toxin
RL I (ApxI) operon.";
RL Infect. Immun. 61:3688-3695(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=S 4074 / Serotype 1;
RX Chang Y., Wang Y., Chin N.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: One of the virulence factors of A.pleuropneumoniae,
CC which has a strong hemolytic activity and is cytotoxic for
CC alveolar macrophages and neutrophils.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: The Gly-rich region is probably involved in binding
CC calcium, which is required for target cell-binding or cytolytic
CC activity.
CC -!- DOMAIN: The three transmembrane domains are believed to be
CC involved in pore formation by the cytotoxin (By similarity).
CC -!- PTM: Palmitoylated by apxIC. The toxin only becomes active when
CC modified (By similarity).
CC -!- MISCELLANEOUS: ApxIA is partially deleted in serotypes 2, 4, 6, 7,
CC 8, 12, and totally deleted in serotype 3.
CC -!- MISCELLANEOUS: The sequence shown is that of serotype 1.
CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52899; CAA37081.1; -
DR EMBL; X68595; CAA48586.1; -
DR EMBL; X73117; CAA51548.1; -
DR EMBL; U05042; AAB05034.1; -
DR PIR; I39643; I39643.
DR InterPro; IPR001343; Hemlyen_Ca_bind.
DR InterPro; IPR003995; RTX_A.
DR Pfam; PF00353; hemolysinCbind; 6.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CAENDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCIUM; 2.
DR Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW Lipoprotein; Palmitate.
FT TRANSMEM 226 256
FT TRANSMEM 297 326
FT TRANSMEM 367 406
FT TRANSMEM 722 845
FT DOMAIN 722 845
FT REPEAT 722 727
FT REPEAT 731 736
FT REPEAT 740 745
FT REPEAT 749 754
FT REPEAT 758 763
FT REPEAT 767 772
FT REPEAT 776 781
FT REPEAT 785 790
FT REPEAT 794 799
FT REPEAT 813 818

```

```

FT REPEAT 822 827
FT REPEAT 831 836
FT REPEAT 840 845
FT CONFLICT 210 217
FT CONFLICT 374 374
FT CONFLICT 562 562
FT CONFLICT 687 688
SQ SEQUENCE 1023 AA; 110193 MW; F99A88CFC9F1A598 CRC64;

Query Match 1.2%; Score 11; DB 1; Length 1023;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 AQRVAAGLSTT 276
Db 289 AQRVAAGLSTT 299

RESULT 8
RT12 ACTPL STANDARD; PRT; 1023 AA.
ID RT12_ACTPL STANDARD; PRT; 1023 AA.
AC P55129;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RTX-I toxin determinant A from serotypes 5/10 (APX-IA) (Hemolysin IA)
DE (HLY-IA) (Cytolysin IA) (CLY-IA).
GN APXIA OR CLYIA OR HLYIA.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=715;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13039 / Serotype 10;
RX MEDLINE=94276858; PubMed=8007819;
RA Nagai S., Yagihashi T., Ishihama A.;
RT "DNA sequence analysis of an allelic variant of the Actinobacillus
RL pleuropneumoniae-RTX-toxin I (ApxIA) from serotype 10.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K17 / Serotype 5;
RX MEDLINE=96401417; PubMed=8807793;
RA Chin N., Frey J., Chang C.F., Chang Y.F.;
RT "Identification of a locus involved in the utilization of iron by
RL Actinobacillus pleuropneumoniae.";
RL FEMS Microbiol. Lett. 143:1-6(1996).
RN [3]
RP SEQUENCE OF 886-1023 FROM N.A.
RC STRAIN=K17 / Serotype 5;
RX MEDLINE=93366425; PubMed=8359891;
RA Jansen R., Briaire J., Kamp E.M., Gielkens A.L.J., Smits M.A.;
RT "Structural analysis of the Actinobacillus pleuropneumoniae-RTX-toxin
RL I (ApxI) operon.";
RL Infect. Immun. 61:3688-3695(1993).
CC -!- FUNCTION: One of the virulence factors of A.pleuropneumoniae,
CC which has a strong hemolytic activity and is cytotoxic for
CC alveolar macrophages and neutrophils.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: The Gly-rich region is probably involved in binding
CC calcium, which is required for target cell-binding or cytolytic
CC activity.
CC -!- DOMAIN: The three transmembrane domains are believed to be
CC involved in pore formation by the cytotoxin (By similarity).
CC -!- PTM: PALMITOYLATED BY APXIC. The toxin only becomes active when
CC modified (By similarity).
CC -!- MISCELLANEOUS: ApxIA is partially deleted in serotypes 2, 4, 6, 7,
CC 8, 12, and totally deleted in serotype 3.
CC -!- MISCELLANEOUS: The sequence shown is that of serotype 10.
CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```



```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype A1 / PHL101;
RX MEDLINE=89210283; PubMed=2707120;
RA Highlander S.K., Chidambaram M., Engler M.J., Weinstein G.M.;
RT "DNA sequence of the Pasteurella haemolytica leukotoxin gene
RL cluster.";
RL DNA 8.15-28(1989).
RN [3]
RP SEQUENCE OF 884-953 FROM N.A.
RC STRAIN=Serotype A1 / PHL101;
RX MEDLINE=90236888; PubMed=2185213;
RA Highlander S.K., Engler M.J., Weinstein G.M.;
RT "Secretion and expression of the Pasteurella haemolytica Leukotoxin.";
RL J. Bacteriol. 172:2343-2350(1990).
CC -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC cell membranes and cause cell rupture by mechanisms not clearly
CC defined.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: The Gly-rich region is probably involved in binding
CC calcium, which is required for target cell-binding or cytolytic
CC activity.
CC -!- DOMAIN: The three transmembrane domains are believed to be
CC involved in pore formation by the cytotoxin (By similarity).
CC -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC modified (By similarity).
CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M20730; AAA25529.1; -.
DR EMBL; M24197; AAA25543.1; -.
DR PIR; B30169; B30169.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR Pfam; PF00353; hemolysinCabind; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 4.
KW Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW Lipoprotein; Palmitate.
FT TRANSMEM 229 249 POTENTIAL.
FT TRANSMEM 297 318 POTENTIAL.
FT TRANSMEM 366 390 POTENTIAL.
FT DOMAIN 734 784 6 X REPEATS, GLY-RICH (BY SIMILARITY).
FT REPEAT 734 739 1.
FT REPEAT 743 748 2.
FT REPEAT 752 757 3.
FT REPEAT 761 766 4.
FT REPEAT 770 775 5.
FT REPEAT 779 784 6.
FT REPEAT 799 814 FEHVAN -> LSTLQI (IN REF. 2).
FT CONFLICT 409 414 D -> Y (IN REF. 2).
FT CONFLICT 742 742
SQ SEQUENCE 953 AA; 101996 MW; 7F93D113A118C05F CRC64;

```

```

Query Match 1.1%; Score 10; DB 1; Length 953;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 265 LAQRVAAGLS 274
Db 283 LAQRVAAGLS 292

```

```

RESULT 11
LKA3_PASHA

```

```

ID LKA3_PASHA STANDARD; PRT; 953 AA.
AC P55116;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Leukotoxin from serotype T3.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype T3;
RX MEDLINE=9401617; PubMed=8225575;
RA Burrows L.L., Lo R.Y.C., Olah-Winfield E.;
RT "Molecular analysis of the leukotoxin determinants from Pasteurella
RL haemolytica serotypes 1 to 16.";
RL Infect. Immun. 61:5001-5007(1993).
CC -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC cell membranes and cause cell rupture by mechanisms not clearly
CC defined.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: The Gly-rich region is probably involved in binding
CC calcium, which is required for target cell-binding or cytolytic
CC activity.
CC -!- DOMAIN: The three transmembrane domains are believed to be
CC involved in pore formation by the cytotoxin (By similarity).
CC -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC modified (By similarity).
CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U01216; AAB36691.1; -.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabind; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 2.
KW Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW Lipoprotein; Palmitate.
FT TRANSMEM 229 249 POTENTIAL.
FT TRANSMEM 297 318 POTENTIAL.
FT TRANSMEM 381 401 POTENTIAL.
FT DOMAIN 734 784 6 X REPEATS, GLY-RICH.
FT REPEAT 734 739 1.
FT REPEAT 743 748 2.
FT REPEAT 752 757 3.
FT REPEAT 761 766 4.
FT REPEAT 770 775 5.
FT REPEAT 779 784 6.
SQ SEQUENCE 953 AA; 101948 MW; FDBDCE2FDC85FDF2 CRC64;

```

```

Query Match 1.1%; Score 10; DB 1; Length 953;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 443 AERVIAITQQ 452
Db 461 AERVIAITQQ 470

```

```

RESULT 12
LKA3_PASHA

```

ID: LKAB_PASHA STANDARD; PRT; 953 AA.
 AC P55116;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Leukotoxin from serotype A11.
 GN LKTA.
 OS Pasteurella haemolytica.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Mannheimia.
 OX NCBI_TaxID=75985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Serotype A11;
 RX MEDLINE=94041617; PubMed=8225575;
 RA Burrows L.L., Olah-Winfield E., Lo R.Y.C.;
 RT "Molecular analysis of the leukotoxin determinants from Pasteurella
 haemolytica serotypes 1 to 16."
 RL Infect. Immun. 61:5001-5007(1993).
 CC -1- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
 cell membranes and cause cell rupture by mechanisms not clearly
 defined.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DOMAIN: The Gly-rich region is probably involved in binding
 calcium, which is required for target cell-binding or cytolytic
 activity.
 CC -1- DOMAIN: The three transmembrane domains are believed to be
 involved in pore formation by the cytotoxin (BY SIMILARITY).
 CC -1- PTM: Palmitoylated by lktC. The toxin only becomes active when
 modified (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U01215; AAB36689.1; -.
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR InterPro; IPR003995; RtxA.
 DR Pfam; PF00353; hemolysinCbind; 5.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PR00313; CABNDNGRPT.
 DR PRINTS; PR01488; RTXTOXINA.
 DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 4.
 KW Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
 KW Lipoprotein; Palmitate.
 FT TRANSMEM 230 250. POTENTIAL.
 FT TRANSMEM 297 317. POTENTIAL.
 FT TRANSMEM 381 401. POTENTIAL.
 FT DOMAIN 734 784. 6 X REPEATS, GLY-RICH.
 FT REPEAT 734 739. 1.
 FT REPEAT 743 748. 2.
 FT REPEAT 752 757. 3.
 FT REPEAT 761 766. 4.
 FT REPEAT 770 775. 5.
 FT REPEAT 779 784. 6.
 SQ SEQUENCE 953 AA; 102206 MW; 927FF56CFC884F12 CRC64;

Query Match 1.1%; Score 10; DB 1; Length 953;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQRVAAGLS 274
 Db 283 LAQRVAAGLS 292
 |||||

RESULT 13
 LKAA_PASHA

ID: LKAA_PASHA STANDARD; PRT; 955 AA.
 AC P55117;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Leukotoxin from serotype T10.
 GN LKTA.
 OS Pasteurella haemolytica.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Mannheimia.
 OX NCBI_TaxID=75985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Serotype T10;
 RX MEDLINE=96425875; PubMed=8828217;
 RA Lainson F.A., Murray J., Davies R.C., Donachie W.;
 RT "Characterization of epitopes involved in the neutralization of
 Pasteurella haemolytica serotype A1 leukotoxin."
 RL Microbiology 142:2499-2507(1996).
 CC -1- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
 cell membranes and cause cell rupture by mechanisms not clearly
 defined.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DOMAIN: The Gly-rich region is probably involved in binding
 calcium, which is required for target cell-binding or cytolytic
 activity.
 CC -1- DOMAIN: The three transmembrane domains are believed to be
 involved in pore formation by the cytotoxin (BY SIMILARITY).
 CC -1- PTM: Palmitoylated by lktC. The toxin only becomes active when
 modified (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; Z26247; CAA81206.1; -.
 DR PIR; S37145; A35254.
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR InterPro; IPR003995; RtxA.
 DR Pfam; PF00353; hemolysinCbind; 5.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PR00313; CABNDNGRPT.
 DR PRINTS; PR01488; RTXTOXINA.
 DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 4.
 KW Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
 KW Lipoprotein; Palmitate.
 FT TRANSMEM 299 319. POTENTIAL.
 FT TRANSMEM 361 381. POTENTIAL.
 FT TRANSMEM 383 403. POTENTIAL.
 FT DOMAIN 736 786. 6 X REPEATS, GLY-RICH.
 FT REPEAT 736 741. 1.
 FT REPEAT 745 750. 2.
 FT REPEAT 754 759. 3.
 FT REPEAT 763 768. 4.
 FT REPEAT 772 777. 5.
 FT REPEAT 781 786. 6.
 SQ SEQUENCE 955 AA; 102187 MW; B60F2DB8168EBCAF CRC64;

Query Match 1.1%; Score 10; DB 1; Length 955;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQRVAAGLS 274
 Db 285 LAQRVAAGLS 294
 |||||

RESULT 14

YF14 MYCTU
ID YF14 MYCTU STANDARD; PRT; 262 AA.
AC P71793;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative glycosyl transferase RV1514C (SC 2.---).
GN RV1514C OR M1564 OR M1564 OR M1564.36C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaita F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayar L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z79701; CAB02028.1; --
CC EMBL; AE007023; AAK45831.1; --
CC F1R; E70714; E70714.
CC TIGR; M1564; --
CC Tuberculin; RV1514C; --
CC InterPro; IPR001173; Glyco trans 2.
CC Pfam; PF00535; Glycosyl transferase; 1.
KW Hypothetical protein; Transferase; Glycosyltransferase;
KW Complete proteome.
SQ SEQUENCE 262 AA; 28965 MW; 6B29BF8D31923E75 CRC64;
Query Match 0.9%; Score 8; DB 1; Length 262;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 746 DGGSGDDV 753
DB 42 DGGSGDDV 49
|||||
RESULT 15
LIPA_CAUCR
ID LIPA_CAUCR STANDARD; PRT; 325 AA.
AC Q9A718;
DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lipic acid synthetase (Lip-syn) (Lipoate synthase).
GN LIPA OR CC1735.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CH15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.C., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -1- FUNCTION: Synthesis of alpha-(+)-lipoic acid. It may be involved
CC in the sulfur insertion chemistry in lipoate biosynthesis (By
CC similarity).
CC -1- PATHWAY: Lipoate biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE BIOTIN AND LIPOIC ACID SYNTHETASES
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE005847; AAK23711.1; --
CC F1R; C87464; C87464.
CC TIGR; CC1735; --
CC HAMAP; MF 00206; --; 1.
CC InterPro; IPR006638; Elp3.
CC InterPro; IPR003698; Lipoate synth.
CC Pfam; PF04055; Radical SAM; 1.
CC SMART; SM00729; Elp3; 1.
CC TIGRFAWS; TIGR00510; lpaA; 1.
KW Iron-sulfur; Complete proteome.
FT METAL 90 90 IRON-SULFUR (POTENTIAL).
FT METAL 94 94 IRON-SULFUR (POTENTIAL).
FT METAL 97 97 IRON-SULFUR (POTENTIAL).
SQ SEQUENCE 325 AA; 36161 MW; 2A1606CD9C3B6400 CRC64;
Query Match 0.9%; Score 8; DB 1; Length 325;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 715 DDLDDGGA 722
DB 135 DDLDDGGA 142
|||||
RESULT 16
Y384_MYCGE
ID Y384_MYCGE STANDARD; PRT; 433 AA.
AC P47624;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable GTP-binding protein MG384.
GN MG384.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;


```

[1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=ATCC 33530 / G-37;
RX  MEDLINE=96026346; PubMed=7569993;
RA  Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA  Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA  Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA  Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA  Tomb J.-F., Dougherty B.A., Bost K.F., Hu P.-C., Lucier T.S.,
RA  Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT  "The minimal gene complement of Mycoplasma genitalium.";
RL  Science 270:397-403(1995).
CC  -1- SIMILARITY: Belongs to the GTP1 / OBG family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U39720; AAC71611.1; -
DR  PIR; E64242; E64242.
DR  TIGR; MG384; -.
DR  InterPro; IPR006074; GTP1/OBG_dom.
DR  InterPro; IPR006073; GTP1_OBG.
DR  InterPro; IPR006169; GTP1_OBG_sub.
DR  InterPro; IPR005225; Small GTP.
DR  Pfam; PF01018; GTP1_OBG; 1.
DR  PRINTS; PR00326; GTP1_OBG.
DR  TIGRPFAMS; TIGR00231; small GTP; 1.
DR  PROSITE; PS00905; GTP1_OBG; 1.
KW  Hypothetical protein; GTP-binding; Complete proteome.
FT  NP_BIND 166 173 GTP (BY SIMILARITY).
FT  NP_BIND 212 216 GTP (BY SIMILARITY).
FT  NP_BIND 282 285 GTP (BY SIMILARITY).
SQ  SEQUENCE 433 AA; 48166 MW; 08BFC7BC794BC3BE CRC64;

Query Match 0.9%; Score 8; DB 1; Length 433;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  90 KLEKFLOK 97
    |||||
Db  295 KLEKFLOK 302

RESULT 17
EGLC_RHIME
ID _EGLC_RHIME STANDARD; PRT; 465 AA.
AC Q9Z3Q2;
DT 30-MAY-2000 (Rel. 39, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Endo-1,3-4-beta-galactanase egLC (EC 3.2.1.-) (Succinoglycan
DE biosynthesis protein egLC).
GN EGLC OR RA0864 OR SWA1587.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=CXMI-105;
RX  MEDLINE=99413305; PubMed=10485295;
RA  Sharypova L.A., Yurgel S.N., Keller M., Simarov B.V., Puehler A.,
RA  Becker A.;
RT  "The eff-482 locus of Sinorhizobium meliloti CXMI-105 that influences
RT symbiotic effectiveness consists of three genes encoding an
RT endoglycanase, a transcriptional regulator and an adenylate cyclase.";
RL  Mol. Gen. Genet. 261:1032-1044(1999).

```

```

[2]
RN  SEQUENCE FROM N.A.
RP  STRAIN=1021;
RX  MEDLINE=21396509; PubMed=11481432;
RA  Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA  Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA  Gural M., Hong A., Huizar L., Hyman R.W., Kahn M.L.,
RA  Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA  Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT  "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSyma megaplasmid.";
RL  Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
CC  -1- FUNCTION: CLEAVES HIGH MOLECULAR WEIGHT SUCCINOGLYCAN TO YIELD LMW
CC  SUCCINOGLYCAN. DYNAMICALLY REGULATES THE MOLECULAR WEIGHT
CC  DISTRIBUTION OF SUCCINOGLYCAN BY CLEAVING NASCENT SUCCINOGLYCAN
CC  ONLY DURING A LIMITED PERIOD AFTER ITS SYNTHESIS, PERHAPS BEFORE
CC  IT UNDERGOES A TIME-DEPENDENT CHANGE IN ITS CONFORMATION OR
CC  AGGREGATION STATE (BY SIMILARITY).
CC  -1- PATHWAY: Exopolysaccharide biosynthesis.
CC  -1- SUBCELLULAR LOCATION: SECRETED. PROBABLY BY A TYPE-III SECRETION
CC  SYSTEM (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AJ225896; CAB38101.1; -
DR  EMBL; AB007273; AAK65522.1; -
DR  PIR; H95369; H95369.
DR  InterPro; IPR000757; Glyco_hydro_16.
DR  InterPro; IPR001343; Hemlysn Ca Bind.
DR  Pfam; PF00722; Glyco_hydro_16; 1.
DR  Pfam; PF00353; hemolysincabind; 3.
DR  PRINTS; PR00313; CABDNGRPT.
DR  PROSITE; PS01034; GLYCOSYL_HYDROL_F16; FALSE NEG.
DR  PROSITE; PS00330; HEMOLYSIN_CALCIIUM; FALSE NEG.
KW  Exopolysaccharide synthesis; Glycosidase; Hydrolase; Plasmid;
KW  Complete proteome.
FT  DOMAIN 275 465 CATALYTIC
FT  ACT_SITE 349 349 NUCLEOPHILE (BY SIMILARITY).
FT  ACT_SITE 354 354 PROTON DONOR (BY SIMILARITY).
FT  CONFLICT 52 52 I -> T (IN REF. 1).
SQ  SEQUENCE 465 AA; 49614 MW; 12CB879AED9E6558 CRC64;

Query Match 0.9%; Score 8; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  755 NGGAGNDV 762
    |||||
Db  123 NGGAGNDV 130

RESULT 18
MAK5 YEAST
ID _MAK5 YEAST STANDARD; PRT; 773 AA.
AC P38112;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE ATP-dependent RNA helicase MAK5.
GN MAK5 OR YBR142W OR YBR1119.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RN  SEQUENCE FROM N.A.

```

```

RC STRAIN=S288C;
RX MEDLINE=95274325; PubMed=7754712;
RA Zgulinski M., Becam A.-M., Grzybowska E., Lacroute F., Migdalski A.,
RA Stoninski P.P., Sokolowska B., Herbert C.J.;
RT "The sequence of 12.5 kb from the right arm of chromosome II predicts
RT a new N-terminal sequence for the IRA1 protein and reveals two new
RT genes, one of which is a DEAD-box helicase.";
RL Yeast 10:1227-1234 (1994).
RN [2]
RP SEQUENCE OF 770-773 FROM N.A.
RC STRAIN=S288C;
RA Etian K.-D., Koetter P., Rose M., Becker J., Grey M., Li Z.,
RA Niesenmann E., Schenk-Groeninger R., Servos J., Wehner E.,
RA Wolter R., Brendel M., Bauer J., Braun H., Dern K., Duesterhus S.,
RA Gruenbein R., Hedges D., Klesau P., Korol S., Krebs B., Proft M.,
RA Siegers K., Baur A., Boles E., Miosga T.,
RA Schaff-Gerstenschlaeger I., Zimmermann F.K.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC 1- FUNCTION: INVOLVED IN MAINTENANCE OF DSRNA KILLER PLASMID.
CC 1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z36011; CAA85100.1; -.
DR EMBL; X78937; CAA55539.1; -.
DR PIR; S46011; S46011.
DR HSSP; Q58083; 1HV8.
DR SGD; S0000346; MAK5.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000629; DEAD box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW Helicase; ATP-binding.
FT NP_BIND 215 222 ATP (POTENTIAL).
FT SITE 333 336 DEAD BOX.
SQ SEQUENCE 773 AA; 87048 MW; C4FF2FB5804FFB9F CRC64;

Query Match 0.9%; Score 8; DB 1; Length 773;
Best Local Similarity 100.0%; Pred.No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 KAADELGI 53
Db 667 KAADELGI 674

RESULT 19
ID CHEA BORBU STANDARD; PRT; 864 AA.
AC Q4737; P70857; Q4877;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chemotaxis protein cheA (EC 2.7.3.-).
GN CHEA OR BB0669.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=212;
RX MEDLINE=98438936; PubMed=9765799;
RA Trueba G.A., Old I.G., Saint-Girons I., Johnson R.C.;

```

```

RT "A cheA cheW operon in Borrelia burgdorferi, the agent of Lyme
RT disease.";
RL Res. Microbiol. 148:191-200 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang M.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Peterson J., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Ueberback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586 (1997).
RN [3]
RP SEQUENCE OF 1-30 FROM N.A.
RC STRAIN=212;
RX MEDLINE=97144545; PubMed=8990312;
RA Ge Y., Charon N.W.;
RT "An unexpected flaA homolog is present and expressed in Borrelia
RT burgdorferi.";
RL J. Bacteriol. 179:552-556 (1997).
CC 1- FUNCTION: INVOLVED IN THE TRANSMISSION OF SENSORY SIGNALS FROM
CC THE CHEMORECEPTORS TO THE FLAGELLAR MOTORS. CHEA IS
CC AUTOPHOSPHORYLATED; IT CAN TRANSFER ITS PHOSPHATE GROUP TO EITHER
CC CHEB OR CHEY (BY SIMILARITY).
CC 1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC 1- SIMILARITY: Contains 1 cheW-like domain.
CC 1- SIMILARITY: Contains 1 histidine kinase domain.
CC 1- SIMILARITY: Contains 1 HPT domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U28962; AAB96835.1; -.
DR EMBL; AE001168; AAC67024.1; -.
DR EMBL; U62900; AAC44771.1; -.
DR EMBL; X91907; CAA63002.1; -.
DR PIR; D70183; D70183.
DR HSSP; Q56310; 1B3Q.
DR TIGR; BB0669; -.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR002545; CheW.
DR InterPro; IPR004105; H-kinase dim.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR002570; Hpt_.
DR Pfam; PF01584; CheW; 1.
DR Pfam; PF02895; H-kinase_dim; 1.
DR Pfam; PF02518; HATPase_C; 1.
DR Pfam; PF01627; Hpt; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR ProDom; PD003142; Hpt_N; 1.
DR SMART; SM00360; CheW; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00073; HPT; 1.
DR PROSITE; PS50851; CHEW; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50894; HPT; 1.
KW Sensory transduction; Transferrase; Kinase; Phosphorylation;
KW Chemotaxis; Complete proteome.
FT DOMAIN 1 108 HPT.
FT DOMAIN 480 725 HISTIDINE KINASE.
FT DOMAIN 727 864 CHEW-LIKE.

```

FT MOD RES 51 51 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CONFLICT 11 11 L -> I (IN REF. 1).
 FT CONFLICT 182 182 S -> H (IN REF. 1).
 FT CONFLICT 187 187 S -> G (IN REF. 1).
 FT CONFLICT 239 239 L -> S (IN REF. 1).
 FT CONFLICT 362 362 A -> S (IN REF. 1).
 FT CONFLICT 559 559 S -> P (IN REF. 1).
 SQ SEQUENCE 864 AA; 98352 MW; C1111DE087BE624 CRC64;
 Query Match 0.9%; Score 8; DB 1; Length 864;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 869 ELKKLADE 876
 Db 255 ELKKLADE 262
 RESULT 20
 LKTA ACTAC
 ID LKTA ACTAC STANDARD; PRT; 1050 AA.
 AC P16462;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Leukotoxin.
 GN LKTA OR LTA.
 OS Actinobacillus actinomycetemcomitans (Haemophilus
 actinomycetemcomitans).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Actinobacillus.
 OX NCBI_TaxID=714;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=JP2;
 RX MEDLINE=89359382; PubMed=2670940;
 RA Lally E.T., Golub E.E., Kieba I.R., Taichman N.S., Rosenbloom J.,
 RA Rosenbloom J.C., Gibson C.W., Demuth D.R.;
 RT "Analysis of the Actinobacillus actinomycetemcomitans leukotoxin
 gene. Delineation of unique features and comparison to homologous
 toxins.";
 RT J. Biol. Chem. 264:15451-15456(1989).
 RL J. Biol. Chem. 264:15451-15456(1989).
 CC -1- FUNCTION: One of the virulence factors of A.actinomycetemcomitans
 might be a cytotoxin, possibly the membrane-bound hemolysin. (By
 CC -1- SUBCELLULAR LOCATION: Outer-membrane associated or secreted (By
 CC similarity).
 CC -1- DOMAIN: The Gly-rich region is probably involved in binding
 CC calcium, which is required for target cell-binding or cytolytic
 CC activity.
 CC -1- DOMAIN: The three transmembrane domains are believed to be
 CC involved in pore formation by the cytotoxin (BY SIMILARITY).
 CC -1- PTM: Palmitoylated by lktC. The toxin only becomes active when
 CC modified (By similarity).
 CC -1- MISCELLANEOUS: Its target cell specificity is restricted to human
 CC and some non-human cells of the mononuclear lineage.
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M27399; AAA21922.1; -;
 DR InterPro; IPR001343; Hemlyan_Ca_bind.
 DR InterPro; IPR003995; RtxA.
 DR Pfam; PF00353; hemolysinCbind; 7.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PR00313; CAENDNGRPT.
 DR PRINTS; PR01488; RTXTOXINA.
 DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 5.

KW Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
 KW Lipoprotein; Palmitate.
 FT TRANSMEM 339 359 POTENTIAL.
 FT TRANSMEM 408 429 POTENTIAL.
 FT TRANSMEM 477 501 POTENTIAL.
 FT DOMAIN 722 844 14 X REPEATS, GLY-RICH.
 FT REPEAT 722 727 1.
 FT REPEAT 731 736 2.
 FT REPEAT 740 745 3.
 FT REPEAT 749 754 4.
 FT REPEAT 758 763 5.
 FT REPEAT 767 772 6.
 FT REPEAT 776 781 7.
 FT REPEAT 785 790 8.
 FT REPEAT 794 799 9.
 FT REPEAT 803 808 10.
 FT REPEAT 812 817 11.
 FT REPEAT 821 826 12.
 FT REPEAT 830 835 13.
 FT REPEAT 839 844 14.
 SQ SEQUENCE 1050 AA; 114194 MW; 38DF9AA24649F662 CRC64;
 Query Match 0.9%; Score 8; DB 1; Length 1050;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 270 AAGLSTTG 277
 Db 295 AAGLSTTG 302
 RESULT 21
 CYAA BORER
 ID CYAA BORER STANDARD; PRT; 1705 AA.
 AC Q57506; O05179;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Bifunctional hemolysin-adenylate cyclase precursor (Cyclolysin) (ACT)
 DE (AC-HLY) [Contains: Calmodulin-sensitive adenylate cyclase
 DE (SC 4.6.1.1) (ATP pyrophosphatase-lyase) (Adenylyl cyclase); Hemolysin).
 GN CYA OR CYAA.
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=518;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=CIP 9.73;
 RX MEDLINE=96009899; PubMed=7557410;
 RA Betson F., Sismeiro O., Danchin A., Guiso N.;
 RT "Cloning and sequence of the Bordetella bronchiseptica adenylate
 RT cyclase-hemolysin-encoding gene: comparison with the Bordetella
 RT pertussis gene.";
 RN Gene 162:165-166(1995).
 RN [2]
 RP REVISIONS TO 1555-1558.
 RA Danchin A., Boursaux-Eude C.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS ADENYLATE CYCLASE BELONGS TO A SPECIAL CLASS OF
 CC BACTERIAL TOXIN. IT CAUSES WHOOPING COUGH BY ACTING ON MAMMALIAN
 CC CELLS BY ELEVATING CAMP-CONCENTRATION AND THUS DISRUPTS NORMAL
 CC CELL FUNCTION.
 CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
 CC -1- ENZYME REGULATION: ACTIVATED BY HOST CALMODULIN.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
 CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
 CC ACTIVITY (BY SIMILARITY).
 CC -1- PTM: RELEASED IN A PROCESSED FORM.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE ADENYLYL
 CC CYCLASE CLASS-2 FAMILY.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE RTX

```
CC ----- PROKARYOTIC TOXIN FAMILY. -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z37112; CAA85481.2; -.
DR HSP; P40136; 1K90.
DR InterPro; IPR005165; Anthrax_toxA.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF03497; Anthrax_toxA; 1.
DR Pfam; PF00353; hemolysinCbind; 17.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CAENDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN CALCIUM; 5.
KW Lyase; CAMP biosynthesis; ATP-binding; Hemolysis; Toxin; Virulence;
KW Whoooping cough; Calcium-binding; Repeat; Lipoprotein; Palmitate.
FT CHAIN 1 312
FT CYCLASE.
FT HEMOLYSIN (BY SIMILARITY TO E.COLI
FT HEMOLYSIN HLIA).
FT A, CATALYTIC.
FT B, ALA/GLY-RICH.
FT C.
FT D, ASP/GLY-RICH.
FT NP BIND 349 356
FT DOMAIN 912 1609
FT REPEAT 912 917 1.
FT REPEAT 1014 1019 2.
FT REPEAT 1023 1028 3.
FT REPEAT 1032 1037 3.
FT REPEAT 1041 1046 4.
FT REPEAT 1050 1055 5.
FT REPEAT 1059 1064 6.
FT REPEAT 1079 1084 7.
FT REPEAT 1164 1169 8.
FT REPEAT 1173 1178 9.
FT REPEAT 1182 1187 10.
FT REPEAT 1202 1207 11.
FT REPEAT 1279 1284 12.
FT REPEAT 1288 1293 13.
FT REPEAT 1297 1302 14.
FT REPEAT 1306 1311 15.
FT REPEAT 1315 1320 16.
FT REPEAT 1324 1329 17.
FT REPEAT 1344 1349 18.
FT REPEAT 1420 1425 19.
FT REPEAT 1429 1434 20.
FT REPEAT 1438 1443 21.
FT REPEAT 1447 1452 22.
FT REPEAT 1555 1560 23.
FT REPEAT 1564 1569 24.
FT REPEAT 1573 1578 25.
FT REPEAT 1582 1587 26.
FT REPEAT 1592 1597 27.
FT REPEAT 1604 1609 28.
FT LIPID 859 859
FT PALMITATE (BY SIMILARITY).
FT PALMITATE (BY SIMILARITY).
FT LIPID 982 982
SQ SEQUENCE 1705 AA; 177153 MW; C43B30F586C835A CRC64;
Query Match 0.9%; Score 8; DB 1; Length 1705;
Best Local Similarity 100.0%; Pred.No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 720 GGAGDRL 727
Db 1032 GGAGDRL 1039
-----
RESULT 22
RL24 PROVV STANDARD; PRT; 18 AA.
ID RL24 PROVV STANDARD; PRT; 18 AA.
AC P20032;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 50S ribosomal protein L24 (Fragment).
GN RPLX.
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Proteus.
OX NCBI_TaxID=585;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89125589; PubMed=2464692;
RA Cerretti D.P., Mattheakis L.C., Kearney K.R., Vu L., Nomura M.;
RT "Translational regulation of the spc operon in Escherichia coli.
RT Identification and structural analysis of the target site for S8
RT repressor protein.";
RL J. Mol. Biol. 204:309-329(1988).
CC -!- FUNCTION: THIS PROTEIN IS FOUND IN THE RIBONUCLEOPROTEIN CORE AND
CC IS INVOLVED IN THE EARLY ASSEMBLY OF THE 50S SUBUNIT. IT IS NOT
CC INVOLVED IN THE FUNCTIONS OF THE MATURE 50S SUBUNIT.
CC -!- SIMILARITY: BELONGS TO THE L24P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M36264; AAA25661.1; -.
DR InterPro; IPR005825; Ribosomal_L24_26.
DR PROSITE; PS01108; RIBOSOMAL_L24; PARTIAL.
KW Ribosomal protein.
FT NON TER 1 1
FT SEQUENCE 18 AA; 2202 MW; 43606F53C5CBA57C CRC64;
Query Match 0.8%; Score 7; DB 1; Length 18;
Best Local Similarity 100.0%; Pred.No. 8.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 481 FEDGKKV 487
Db 1 FEDGKKV 7
-----
RESULT 23
CECA BOMMO STANDARD; PRT; 63 AA.
ID CECA BOMMO STANDARD; PRT; 63 AA.
AC Q27239;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cecropin A precursor.
GN CECA.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C108; TISSUE=Larval fat body;
RA MEDLINE=94369101; PubMed=7765280;
RA Yamano Y., Matsumoto M., Inoue K., Kawabata T., Morishima I.;
RT "Cloning of cDNAs for cecropins A and B, and expression of the genes
RT in the silkworm, Bombyx mori.";
-----
Query Match 0.9%; Score 8; DB 1; Length 1705;
Best Local Similarity 100.0%; Pred.No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 720 GGAGDRL 727
Db 1032 GGAGDRL 1039
```

```

RL Biosci. Biotechnol. Biochem. 58:1476-1478(1994).
CC -1- FUNCTION: CECROPINS HAVE LYTIC AND ANTIBACTERIAL ACTIVITY AGAINST
CC SEVERAL GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Highest expression in fat body and hemocytes.
CC Is also expressed in Malpighian tubules and to a much lesser
CC extent in midgut. Not present in silk gland.
CC -1- SIMILARITY: BELONGS TO THE CECROPIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D17394; AAC04217.1; -.
DR EMBL; S74376; AAC60515.1; -.
DR PIR; JC2295; CECROTA.
DR InterPro; IPR000875; Cecropin.
DR InterPro; IPR003254; IIP cecropin.
DR Pfam; PF00272; cecropin; 1.
DR ProDom; PD003996; IIP cecropin; 1.
DR PROSITE; PS00268; CECROPIN; 1.
KW Insect immunity; Antibiotic; Hemolymph; Amidation; Multigene family;
KW Signal.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 26 BY SIMILARITY.
FT CHAIN 27 61 CECROPIN A.
FT MOD_RES 61 61 AMIDATION (G-62 PROVIDE AMIDE GROUP)
FT SEQUENCE 63 AA; 6762 MW; 6A1C39975516D86A CRC64;

Query Match 0.8%; Score 7; DB 1; Length 63;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 347 ALGAVSA 353
DB 16 ALGAVSA 22
|||||

RESULT 24
SCX1_TITSE STANDARD; PRT; 80 AA.
ID SCX1_TITSE
AC P01496; P91788;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Toxin IV-5 precursor (Tityustoxin) (Fragment).
OS Tityus serrulatus (Brazilian scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Butthidea; Butthidae; Tityus.
OX NCBI_TaxID=6887;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96243503; PubMed=8711758;
RA Corona M., Zurita M., Possani L.D., Becerril B.;
RT "Cloning and characterization of the genomic region encoding toxin
RL Toxin IV-5 from the scorpion Tityus serrulatus Lutz and Mello.";
RN [2]
RP SEQUENCE OF 14-75.
RC TISSUE=Venom;
RX MEDLINE=91131623; PubMed=1993690;
RA Possani L.D., Martin B.M., Fletcher M.D., Fletcher P.L. Jr.;
RT "Discharge effect on pancreatic exocrine secretion produced by toxins
RL purified from Tityus serrulatus scorpion venom.";
RN [3]
RP SEQUENCE OF 14-43.
RC TISSUE=Venom;

```

```

RA Possani L.D., Martin B.M., Mochca-Morales J., Svendsen I.;
RT "N-terminal sequence of toxin IV-5 from the venom of the scorpion
RT Tityus serrulatus.";
RL Toxinon 20:75-76(1982).
CC -1- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC of the activated channels, thereby blocking neuronal transmission.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC BETA-TOXIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S82286; AAB37719.2; -.
DR HSSP; P46086; INRB.
DR InterPro; IPR003614; Knot1.
DR InterPro; IPR002061; Scorpion_toxinL.
DR Pfam; PF00537; toxin_3; 1.
DR ProDom; PD000908; Scorpion_toxinL; 1.
DR SMART; SM00505; Knot1; 1.
KW Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW Amidation; Signal.
FT NON_TER 1 1
FT SIGNAL <1 13
FT CHAIN 14 77 TOXIN IV-5.
FT DISULFID 25 75 BY SIMILARITY.
FT DISULFID 29 51 BY SIMILARITY.
FT DISULFID 37 58 BY SIMILARITY.
FT DISULFID 41 60 BY SIMILARITY.
FT MOD_RES 77 77 AMIDATION (G-78 PROVIDE AMIDE GROUP)
FT CONFLICT 70 73 KTNQ -> GST (IN REF. 2).
FT SEQUENCE 80 AA; 9043 MW; 974F46346A435FFD CRC64;

Query Match 0.8%; Score 7; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 521 LITAGTE 527
DB 6 LITAGTE 12
|||||

RESULT 25
SCX4_TITSE STANDARD; PRT; 80 AA.
ID SCX4_TITSE
AC P45659;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Toxins IV/III/V precursor (TSTX-IV) (Tityustoxin IV) (Ts IV)
DE (Fragment).
OS Tityus serrulatus (Brazilian scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Butthidea; Butthidae; Tityus.
OX NCBI_TaxID=6887;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 14-43.
RC TISSUE=Venom;
RX MEDLINE=94192827; PubMed=8143874;
RA Martin-Eauclaire M.-F., Ceard B., Ribeiro A.M., Diniz C.R., Rochat H.;
RA Bougis P.E.;
RT "Biochemical, pharmacological and genomic characterisation of Ts IV,
RL an alpha-toxin from the venom of the South American scorpion Tityus
RL serrulatus.";
RN [3]
RP SEQUENCE OF 14-43.
RC TISSUE=Venom;

```

```
CC of the activated channels, thereby blocking neuronal transmission.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC ALPHA-TOXIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S69808; AAB30413.1; -.
CC PIR; S43674; NTSR4T.
CC HSP; P46066; INRB.
CC InterPro; IPR003614; Koot1.
CC InterPro; IPR002061; Scorpion_toxinL.
CC Pfam; PF00537; toxin_3; 1.
CC ProDom; PD000908; Scorpion_toxinL; 1.
CC SMART; SM00505; Koot1; 1.
CC Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
CC Signal; Amidation.
CC NON_TER 1
CC SIGNAL <1 13
CC CHAIN 14 77
CC CHAIN 14 79
CC CHAIN 14 75
CC DISULFID 25 75
CC DISULFID 29 51
CC DISULFID 37 58
CC DISULFID 41 60
CC MOD_RES 77
CC SEQUENCE 80 AA; 9025 MW; FDE437BEC1335FFC CRC64;
Query Match 0.8%; Score 7; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 521 LLTAGTE 527
DB 6 LLTAGTE 12
RESULT 26
TRAC_RHISN STANDARD; PRT; 102 AA.
AC P55419;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Probable conjugal transfer protein trac.
GN TRAC OR Y4DT. (strain NGR234).
OS Rhizobium sp.
OG Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: STRONG, TO A.TUMEFACIENS TI PLASMID TRAC.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000070; AAB92440.1; -.
CC Conjugation; Plasmid.
CC SEQUENCE 102 AA; 10205 MW; AA378F93DD06C78B CRC64;
Query Match 0.8%; Score 7; DB 1; Length 102;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 356 SAAAVGS 362
DB 86 SAAAVGS 92
RESULT 27
RL24_ECOLI STANDARD; PRT; 103 AA.
AC P22425; P37438;
DT 21-JUN-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L24.
GN RPLX OR B3309 OR Z4679 OR ECS4174 OR STM3429.
OS Escherichia coli, O157:H7, and
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562; 83334; 602;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=83220807; PubMed=6222285;
RA Caretti D.P., Dean D., Davis G.R., Bedwell D.M., Nomura M.;
RT "The spc ribosomal protein operon of Escherichia coli: sequence and
RT cotranscription of the ribosomal protein genes and a protein export
RT gene.";
RL Nucleic Acids Res. 11:2599-2616(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Potlatch E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki M., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
```



```

DR Pfam: PF00467; KOW; 1.
DR ProDom: PD001677; Ribosomal_L24; 1.
DR SMART: SM00739; KOW; 1.
DR TIGRFAMs: TIGR01079; rplX_bact; 1.
DR PROSITE: PS01108; RIBOSOMAL_L24; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 103 AA; 11285 MW; 3AD490337702F19 CRC64;

Query Match 0.8%; Score 7; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 481 FEDGKKV 487
Db 87 FEDGKKV 93

RESULT 29
RL24_PASNU
ID RL24_PASNU STANDARD; PRT; 103 AA.
AC Q9CL41;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L24.
GN RPLX OR RPL24 OR PM1404.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Faustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -!- FUNCTION: THIS PROTEIN IS FOUND IN THE RIBONUCLEOPROTEIN CORE AND
CC IS INVOLVED IN THE EARLY ASSEMBLY OF THE 50S SUBUNIT. IT IS NOT
CC INVOLVED IN THE FUNCTIONS OF THE MATURE 50S SUBUNIT
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE L24P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE006177; AAK03488.1; -.
CC InterPro: IPR005824; KOW.
CC InterPro: IPR006646; KOW sub.
CC InterPro: IPR003256; Ribosomal_L24.
CC InterPro: IPR005825; Ribosomal_L24_26.
CC Pfam: PF00467; KOW; 1.
CC ProDom: PD001677; Ribosomal_L24; 1.
CC SMART: SM00739; KOW; 1.
CC TIGRFAMs: TIGR01079; rplX_bact; 1.
CC PROSITE: PS01108; RIBOSOMAL_L24; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 103 AA; 11239 MW; 0A325EF73B604850 CRC64;

Query Match 0.8%; Score 7; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 481 FEDGKKV 487
Db 87 FEDGKKV 93

RESULT 29
RL24_PASNU
ID RL24_PASNU STANDARD; PRT; 103 AA.
AC Q9CL41;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L24.
GN RPLX OR RPL24 OR PM1404.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Faustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -!- FUNCTION: THIS PROTEIN IS FOUND IN THE RIBONUCLEOPROTEIN CORE AND
CC IS INVOLVED IN THE EARLY ASSEMBLY OF THE 50S SUBUNIT. IT IS NOT
CC INVOLVED IN THE FUNCTIONS OF THE MATURE 50S SUBUNIT
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE L24P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE006177; AAK03488.1; -.
CC InterPro: IPR005824; KOW.
CC InterPro: IPR006646; KOW sub.
CC InterPro: IPR003256; Ribosomal_L24.
CC InterPro: IPR005825; Ribosomal_L24_26.
CC Pfam: PF00467; KOW; 1.
CC ProDom: PD001677; Ribosomal_L24; 1.
CC SMART: SM00739; KOW; 1.
CC TIGRFAMs: TIGR01079; rplX_bact; 1.
CC PROSITE: PS01108; RIBOSOMAL_L24; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 103 AA; 11239 MW; 0A325EF73B604850 CRC64;

Query Match 0.8%; Score 7; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 481 FEDGKKV 487
Db 87 FEDGKKV 93

RESULT 30
RL24_BUCAK
ID RL24_BUCAK STANDARD; PRT; 104 AA.
AC PF6177;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 50S ribosomal protein L24.
GN RPLX.
OS Buchnera aphidicola (subsp. Acyrthosiphon kondoi) (Acyrthosiphon
OS kondoi symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=42474;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KuraShiki;
RX MEDLINE=96051390; PubMed=7584036;
RA Abe R., Yamashita A., Isono K.;
RT "Cloning and characterization of the ribosomal protein genes in the
RT spc operon of a prokaryotic endosymbiont of the pea aphid,
RT Acyrthosiphon kondoi.";
RL DNA Res. 1:103-114(1994).
CC -!- FUNCTION: THIS PROTEIN IS FOUND IN THE RIBONUCLEOPROTEIN CORE AND
CC IS INVOLVED IN THE EARLY ASSEMBLY OF THE 50S SUBUNIT. IT IS NOT
CC INVOLVED IN THE FUNCTIONS OF THE MATURE 50S SUBUNIT
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE L24P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D31786; BAA06586.1; -.
CC InterPro: IPR005824; KOW.
CC InterPro: IPR006646; KOW sub.
CC InterPro: IPR003256; Ribosomal_L24.
CC InterPro: IPR005825; Ribosomal_L24_26.
CC Pfam: PF00467; KOW; 1.
CC ProDom: PD001677; Ribosomal_L24; 1.
CC SMART: SM00739; KOW; 1.
CC TIGRFAMs: TIGR01079; rplX_bact; 1.
CC PROSITE: PS01108; RIBOSOMAL_L24; 1.
KW Ribosomal protein.
SQ SEQUENCE 104 AA; 11336 MW; F3553FE6418BF47C CRC64;

Query Match 0.8%; Score 7; DB 1; Length 104;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 481 FEDGKKV 487
Db 87 FEDGKKV 93

RESULT 31
NU4M_CAICR
ID NU4M_CAICR STANDARD; PRT; 111 AA.
AC Q34076;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3) (Fragment).
GN MTND4 OR ND4.
OS Calman crocodilus (Spectacled calman) (Calman sclerops).
OS Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Calman.
OX NCBI_TaxID=8499;

```



```

RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=96073446; PubMed=7476123;
RA  Kumazawa Y., Nishida M.;
RT  "Variations in mitochondrial tRNA gene organization of reptiles as
RL  phylogenetic markers.";
RL  Mol. Biol. Evol. 12:759-772(1995).
CC  -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; D38190; BAA07383.1; -.
DR  InterPro; IPR001750; Oxidored q1.
DR  Pfam; PF00361; oxidored q1; 1.
KW  Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
FT  NON_TER
SQ  SEQUENCE 111 AA; 12070 MW; D211FBBA07A42D83 CRC64;

Query Match 0.8%; Score 7; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  518 TSPLLTA 524
DB  |||||
   5 TSPLLTA 11

RESULT 32
RL7 NEILA
ID  RL7 NEILA STANDARD; PRT; 122 AA.
AC  Q9ETV2;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  50S ribosomal protein L7/L12.
GN  RPLL.
OS  Neisseria lactamica, and
OS  Neisseria sicca.
OC  Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC  Neisseriaceae; Neisseria.
OX  NCBI_TaxID=486, 490;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  SPECIES=N.lactamica, and N.sicca; STRAIN=NRL 8828, and NRL 30016;
RA  Spence J.M., Clark V.L.;
RT  "Alterations in protein profiles associated with induction of the
RT  contact-induced enhanced invasion phenotype of Neisseria
RT  gonorrhoeae.";
RL  Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC  -1- FUNCTION: SEEMS TO BE THE BINDING SITE FOR SEVERAL OF THE FACTORS
CC  INVOLVED IN PROTEIN SYNTHESIS AND APPEARS TO BE ESSENTIAL FOR
CC  ACCURATE TRANSLATION (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF312972; AAG34164.1; -.
DR  EMBL; AF312974; AAG34166.1; -.
DR  HSSP; P02392; 1CTF.
DR  HAMAP; MF 00368; -; 1.
DR  InterPro; IPR000206; Ribosomal_L12.
```

```

DR  Pfam; PF00542; Ribosomal_L12; 1.
DR  ProDom; PD001326; Ribosomal_L12; 1.
DR  TIGRFAMs; TIGR00855; L12; 1.
KW  Ribosomal protein.
SQ  SEQUENCE 122 AA; 12567 MW; 7A87249AAC0E2D4D CRC64;

Query Match 0.8%; Score 7; DB 1; Length 122;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  354 GVSAAAV 360
DB  |||||
   32 GVSAAAV 38

RESULT 33
RL7 NEIMA
ID  RL7 NEIMA STANDARD; PRT; 122 AA.
AC  P80716;
DT  01-OCT-1996 (Rel. 34, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  50S ribosomal protein L7/L12.
GN  RPLL OR NMA0143 OR NMB0131.
OS  Neisseria meningitidis (serogroup A), and
OS  Neisseria meningitidis (serogroup B).
OC  Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC  Neisseriaceae; Neisseria.
OX  NCBI_TaxID=65699, 491;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=22491 / Serogroup A / Serotype 4A;
RC  MEDLINE=20222556; PubMed=10761919;
RA  Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA  Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA  Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA  Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA  Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA  Whitehead S., Spratt B.G., Barrell B.G.;
RT  "Complete DNA sequence of a serogroup A strain of Neisseria
RT  meningitidis 22491.";
RL  Nature 404:502-506(2000).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=MC58 / Serogroup B;
RC  MEDLINE=20175755; PubMed=10710307;
RA  Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA  Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA  Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA  Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA  Mason T., Ciecko A., Parksey D.S., Blair E., Citron H., Clark E.B.,
RA  Cotton M.D., Uterback T.R., Khouri H., Qin H., Vamathevan J.,
RA  Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA  Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT  "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT  MC58.";
RL  Science 287:1809-1815(2000).
RN  [3]
RP  SEQUENCE OF 1-20.
RC  STRAIN=44/76 / Serogroup B / Serotype 15 / Subtype 16;
RC  MEDLINE=97177772; PubMed=9025278;
RA  Kolberg J., Holby E.A., Lopez R., Sletten K.;
RT  "Monoclonal antibodies against Streptococcus pneumoniae detect
RT  epitopes on eubacterial ribosomal proteins L7/L12 and on
RT  streptococcal elongation factor Ts.";
RL  Microbiology 143:55-61(1997).
CC  -1- FUNCTION: SEEMS TO BE THE BINDING SITE FOR SEVERAL OF THE FACTORS
CC  INVOLVED IN PROTEIN SYNTHESIS AND APPEARS TO BE ESSENTIAL FOR
CC  ACCURATE TRANSLATION (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AL162752; CAB83458.1; -;
DR EMBL; AE002371; AAF40590.1; -;
DR PIR; H81235; H81235.
DR HSSP; P02392; 1CTF.
DR TIGR; NMB0131; -;
DR HAMAP; MF_00368; -; 1.
DR InterPro; IPR000206; Ribosomal_L12.
DR Pfam; PF00542; Ribosomal_L12; 1.
DR ProDom; PD001326; Ribosomal_L12; 1.
DR TIGRFAMs; TIGR00855; L12; 1.
KW Ribosomal protein; Complete proteome.
FT INIT MET 0
SQ SEQUENCE 122 AA; 12491 MW; 3C962C3A11B6F535 CRC64;

Query Match 0.8%; Score 7; DB 1; Length 122;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 354 GVSAAAV 360
|||||
DB 31 GVSAAAV 37

RESULT 34
RL7_YERPE STANDARD; PRT; 122 AA.
AC Q8ZAP4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L7/L12.
GN RPLL OR YPO3748 OR Y0483.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).

CC -!- FUNCTION: Seems to be the binding site for several of the factors
CC involved in protein synthesis and appears to be essential for
CC accurate translation (By similarity).
CC -!- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AJ414158; CAC93216.1; -;
DR EMBL; AE013649; AM84072.1; -;
DR PIR; AD0456; AD0456.
DR HAMAP; MF_00368; -; 1.
DR InterPro; IPR000206; Ribosomal_L12.
DR Pfam; PF00542; Ribosomal_L12; 1.
DR ProDom; PD001326; Ribosomal_L12; 1.
DR TIGRFAMs; TIGR00855; L12; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 122 AA; 12530 MW; 40A0C74D6B77D69D CRC64;

Query Match 0.8%; Score 7; DB 1; Length 122;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 354 GVSAAAV 360
|||||
DB 33 GVSAAAV 39

RESULT 35
RL7_NEIPE STANDARD; PRT; 123 AA.
AC Q9F5M1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L7/L12.
GN RPLL.
OS Neisseria perflava.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=33053;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRL 30015;
RA Spence J.M., Clark V.L.;
RT "Alterations in protein profiles associated with induction of the
RT contact-induced enhanced invasion phenotype of Neisseria
RT gonorrhoeae."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SEEMS TO BE THE BINDING SITE FOR SEVERAL OF THE FACTORS
CC INVOLVED IN PROTEIN SYNTHESIS AND APPEARS TO BE ESSENTIAL FOR
CC ACCURATE TRANSLATION (By similarity).
CC -!- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF312973; AAG34165.1; -;
DR HSSP; P02392; 1CTF. 1
DR HAMAP; MF_00368; -; 1
DR InterPro; IPR000206; Ribosomal_L12.
DR Pfam; PF00542; Ribosomal_L12; 1.
DR ProDom; PD001326; Ribosomal_L12; 1.
DR TIGRFAMs; TIGR00855; L12; 1.
KW Ribosomal protein.
SQ SEQUENCE 123 AA; 12578 MW; 5A6257CB0DE3A435 CRC64;

Query Match 0.8%; Score 7; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 354 GVSAAV 360
Db 32 GVSAAV 38
|||||

RESULT 36
RL7_RALSO STANDARD; PRT; 124 AA.
AC Q8XU27;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE S0S ribosomal protein L7/L12.
GN RFL1 OR RSC3035 OR RS04722.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002)
CC Involved in protein synthesis and appears to be essential for
CC accurate translation (By similarity).
CC -! SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL646073; CAB16744.1; -.
DR HAMAP; MF 00368; -.
DR InterPro; IPR001813; 60s ribosomal.
DR InterPro; IPR002026; Ribosomal L12.
DR Pfam; PF00428; 60s ribosomal; 1.
DR Pfam; PF00542; Ribosomal_L12; 1.
DR ProDom; PD001326; Ribosomal_L12; 1.
DR TIGRFAMs; TIGR00855; L12; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 124 AA; 13520 MW; 271C3CF71F80269E CRC64;

Query Match 0.8%; Score 7; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 354 GVSAAV 360
Db 32 GVSAAV 38
|||||

RESULT 37
H2AV_STRPU STANDARD; PRT; 125 AA.
AC P08931;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Histone H2A variant (Fragment).
GN H2A.F/Z.

Strongylocentrotus purpuratus (Purple sea urchin).
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Echinoidea; Echinacea; Echinoida; Strongylocentrotidae;
Strongylocentrotus.
NCBI_TaxID=7668;
[1]
SEQUENCE FROM N.A.
MEDLINE=87231039; PubMed=2438657;
Ernst S.G., Miller H., Brenner C.A., Nocente-Mcgrath C., Francis S.,
McIsaac R.;
"Characterization of a cDNA clone coding for a sea urchin histone H2A
variant related to the H2A.F/Z histone protein in vertebrates.";
Nucleic Acids Res. 15:4629-4644(1987).
-! FUNCTION: VARIANT HISTONES H2A ARE SYNTHESIZED THROUGHOUT THE
CELL CYCLE AND ARE VERY DIFFERENT FROM CLASSICAL S-PHASE REGULATED
H2A. THE EXACT FUNCTION OF VARIANT HISTONES H2A IS NOT KNOWN.
-! SUBUNIT: The nucleosome is an octamer containing two molecules
each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
bp of DNA.
-! SUBCELLULAR LOCATION: Nuclear.
-! SIMILARITY: Belongs to the histone H2A family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X05547; CAA29061.1; -.
DR PIR; S07392; S07392.
DR InterPro; IPR004822; Histone_core.
DR InterPro; IPR002119; Histone_H2A.
DR Pfam; PF00125; histone; 1.
DR PRINTS; PR00620; HISTONEH2A.
DR ProDom; PD000522; Histone_H2A; 1.
DR SMART; SM00414; H2A; 1.
DR PROSITE; PS00046; HISTONE_H2A; 1.
KW Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding;
KW Multigene family.
FT NON TER 1
FT TER 1
SQ SEQUENCE 125 AA; 13164 MW; F65A17FDF263823F CRC64;

Query Match 0.8%; Score 7; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 ELDSLK 145
Db 93 ELDSLK 99
|||||

RESULT 38
H2AV_CHICK STANDARD; PRT; 127 AA.
AC P02272;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Histone H2A variant.
GN H2AF.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83221493; PubMed=6574451;
RA Harvey R.P., Whiting J.A., Coles L.S., Krieg P.A., Wells J.R.E.;
RT "H2A.F: an extremely variant histone H2A sequence expressed in the
chicken embryo.";

```

```
RL Proc. Natl. Acad. Sci. U.S.A. 80:2819-2823 (1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89160327; PubMed=2493634;
RA Dalton S., Robins A.J., Harvey R.P., Wells J.R.E.;
RT "Transcription from the intron-containing chicken histone H2A.F gene
is not S-phase regulated.";
RL Nucleic Acids Res. 17:1745-1756 (1989).
CC -!- FUNCTION: VARIANT HISTONES H2A ARE SYNTHESIZED THROUGHOUT THE
CELL CYCLE AND ARE VERY DIFFERENT FROM CLASSICAL S-PHASE REGULATED
H2A. THE EXACT FUNCTION OF VARIANT HISTONES H2A IS NOT KNOWN.
CC -!- SUBUNIT: The nucleosome is an octamer containing two molecules
each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
BP of DNA.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE CHICKEN EMBRYO.
CC -!- SIMILARITY: Belongs to the histone H2A family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; V00414; CAA23705.1; -.
DR EMBL; X13894; CAA32094.1; ALT_SEQ.
DR PIR; S03282; HSC2F.
DR InterPro; IPR004822; Histone core.
DR InterPro; IPR002119; Histone_H2A.
DR Pfam; PF00125; histone; 1.
DR PRINTS; PR00620; HISTONEH2A.
DR ProDom; PD000522; Histone_H2A; 1.
DR SMART; SM00414; H2A; 1.
DR PROSITE; PS00046; HISTONE_H2A; 1.
KW Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding;
KW Multigene family; Embryo.
FT INIT MET 0
SQ SEQUENCE 127 AA; 13377 MW; 865AC88F6832F437 CRC64;

Query Match 0.8%; Score 7; DB 1; Length 127;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 139 ELDSLK 145
Db 95 ELDSLK 101
|||||

RESULT 39
ID H2AZ HUMAN STANDARD; PRT; 127 AA.
AC P17317;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Histone H2A.z (H2A/z).
GN H2AFZ OR H2AZ.
OS Homo sapiens (Human).
OS Mus musculus (Mouse).
OS Rattus norvegicus (Rat), and
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606, 10090, 10116, 9913;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human, Rat, and Bovine; TISSUE=Brain;
RX MEDLINE=86143983; PubMed=3344202;
RA Hatch C.L., Bonner W.M.;
RT "Sequence of cDNAs for mammalian H2A.Z, an evolutionarily diverged
but highly conserved basal histone H2A isoprotein species.";
```

```
RL Nucleic Acids Res. 16:1113-1124 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RX SPECIES=Human;
MEDLINE=90368704; PubMed=1697587;
RA Hatch C.L., Bonner W.M.;
RT "The human histone H2A.Z gene. Sequence and regulation.";
RL J. Biol. Chem. 265:15211-15218 (1990).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse; STRAIN=C57BL/6;
RA Rocha D., Carrier A., Anderson E., Botcherby M., Guenet J.-L.,
Jordan B.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: VARIANT HISTONES H2A ARE SYNTHESIZED THROUGHOUT THE
CELL CYCLE AND ARE VERY DIFFERENT FROM CLASSICAL S-PHASE REGULATED
H2A. THE EXACT FUNCTION OF VARIANT HISTONES H2A IS NOT KNOWN.
CC -!- SUBUNIT: The nucleosome is an octamer containing two molecules
each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
BP of DNA.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the histone H2A family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; X52317; CAA36553.1; -.
DR EMBL; M37583; AAA35984.1; -.
DR EMBL; L10138; AAC61625.1; -.
DR EMBL; X52316; CAA36552.1; -.
DR EMBL; X52318; CAA36554.1; -.
DR EMBL; M37584; AAA41329.1; -.
DR EMBL; M37585; AAA30566.1; -.
DR EMBL; U70494; AAB09578.1; -.
DR PIR; A35881; A35881.
DR PIR; S03642; S03642.
DR PIR; S03644; S03644.
DR PDB; 1F66; 27-NOV-00.
DR Genew; HGNC:4741; H2AFZ.
DR MIM; 142763; -.
DR MGD; MGI:1888388; H2afz.
DR GO; GO:0005718; C:nucleosome; NAS.
DR GO; GO:0003677; F:DNA binding activity; NAS.
DR GO; GO:0007001; P:chromosome organization and biogenesis (sen. .); NAS.
DR GO; GO:0006334; P:nucleosome assembly; NAS.
DR InterPro; IPR004822; Histone core.
DR InterPro; IPR002119; Histone_H2A.
DR Pfam; PF00125; histone; 1.
DR PRINTS; PR00620; HISTONEH2A.
DR ProDom; PD000522; Histone_H2A; 1.
DR SMART; SM00414; H2A; 1.
DR PROSITE; PS00046; HISTONE_H2A; 1.
KW Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding;
KW Multigene family; 3D-structure.
FT INIT MET 0
SQ SEQUENCE 127 AA; 13421 MW; 794215381845F35A CRC64;

Query Match 0.8%; Score 7; DB 1; Length 127;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 139 ELDSLK 145
Db 95 ELDSLK 101
|||||

RESULT 40
RL20_BIFLO
```

ID RL20_BIFLO STANDARD; PRT; 127 AA.
AC Q8G4L1;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 50S ribosomal protein L20.
GN RPLT OR BL1367.
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787;
RA Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA Pessi G., Zwaalen M.-C., Desiere F., Bork P., Delley M.,
RA Pridmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
CC -!- FUNCTION: This protein binds directly to 23S ribosomal RNA and is
necessary for the in vitro assembly process of the 50S ribosomal
subunit. It is not involved in the protein synthesizing functions
of that subunit (By similarity).
CC -!- SIMILARITY: BELONGS TO THE L20P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AE014766; AAN25168.1; -.
DR HAMAP; MF_00382; -; 1.
DR InterPro; IPR005813; L20.
DR InterPro; IPR005812; L20_bact_org.
DR Pfam; PF00453; Ribosomal L20; 1.
DR PRINTS; PR00062; RIBOSOMALL20.
DR ProDom; PD002389; L20; 1.
DR TIGRFAMs; TIGR01032; rplT_bact; 1.
DR PROSITE; PS00937; RIBOSOMAL L20; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 127 AA; 14605 MW; 2EF7D5D28B7BDD99 CRC64;

Query Match 0.8%; Score 7; DB 1; Length 127;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 135 LAGIELD 141
Db 85 LAGIELD 91

Search completed: February 17, 2004, 10:20:08
Job time 1:39 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:17:00 ; Search time 45 Seconds
(without alignments)
5315.884 Million cell updates/sec

Title: US-10-069-799-5
Perfect score: 927
Sequence: 1 MSINVIKSNIQAGLNSTKS.....SSNALQPIPTQIGILAPSV 927

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

SPTREMBL 23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archheap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 826 | 89.1 | 927 | Q93G12 | Q93G12 moraxella b |
| 2 | 14 | 1.5 | 998 | Q46716 | Q46716 escherichia |
| 3 | 14 | 1.5 | 998 | Q85101 | Q85101 escherichia |
| 4 | 14 | 1.5 | 998 | P71223 | P71223 escherichia |
| 5 | 14 | 1.5 | 998 | Q47262 | Q47262 escherichia |
| 6 | 14 | 1.5 | 998 | Q47461 | Q47461 escherichia |
| 7 | 14 | 1.5 | 998 | Q91C58 | Q91C58 escherichia |
| 8 | 14 | 1.5 | 1024 | Q8GA40 | Q8GA40 escherichia |
| 9 | 14 | 1.5 | 1024 | Q8G924 | Q8G924 escherichia |
| 10 | 14 | 1.5 | 1024 | Q8FE01 | Q8FE01 escherichia |
| 11 | 14 | 1.5 | 1049 | Q3RCG8 | Q3RCG8 pasteurilla |
| 12 | 14 | 1.5 | 1052 | Q33NP0 | Q33NP0 actinobacil |
| 13 | 12 | 1.3 | 987 | Q8KW29 | Q8KW29 actinobacil |
| 14 | 12 | 1.3 | 987 | Q8KW26 | Q8KW26 actinobacil |
| 15 | 11 | 1.2 | 956 | Q93NP1 | Q93NP1 actinobacil |
| 16 | 11 | 1.2 | 1022 | Q93NP2 | Q93NP2 actinobacil |

Q8MR60 drosophila
Q9ZFH0 azotobacter
Q8RMZ8 azotobacter
Q9EV24 manheimia
Q9EV31 pasteurilla
Q9ETX2 manheimia
Q9EV27 pasteurilla
Q9EV29 pasteurilla
Q9EV25 manheimia
Q9ETG5 pasteurilla
Q9EV23 manheimia
Q9EV33 pasteurilla
Q9EV34 pasteurilla
Q9EV32 pasteurilla
Q9EV30 pasteurilla
Q9EV28 pasteurilla
Q9EV26 manheimia
Q9EUE1 pasteurilla
Q9EUD4 pasteurilla
Q9EV22 pasteurilla
Q44494 azotobacter
Q9VL42 drosophila
Q8YKJ3 anabaena sp
Q9RDH8 streptomyces
Q9R9H2 pseudomonas
Q8PH96 xanthomonas
Q8UH95 agrobacteri
Q8LHL1 oryza sativ
Q8D3W2 vibrio vuln
Q51868 pasteurilla
Q5199 rhizobium m
Q51865 pasteurilla
Q8ZAF8 yersinia pe
Q8PJ41 xanthomonas
Q8UA74 agrobacteri
Q8KE19 chlorobium
Q8Y0F0 ralstonia s
Q8LP4 tulipa bake
Q8U9U5 agrobacteri
Q8ZK75 salmonella
Q8Z156 salmonella
Q44223 anabaena sp
Q8G3J1 bifidobacte
Q8F4G0 leptospira
Q9X5A3 treponema d
Q92VW6 rhizobium m
Q92XC8 rhizobium m
Q92XB8 rhizobium m
Q92FU0 listeria in
Q8YAV0 listeria mo
Q9WV2 agrobacteri
Q8UKU5 agrobacteri
Q9FV87 petroselinu
Q9FV86 petroselinu
Q17900 caenorhabdi
Q92LP8 rhizobium m
Q8ZVD7 pyrobaculum
Q17898 caenorhabdi
Q99KX3 mus musculu
Q9LGS6 oryza sativ
Q8YV5 ralstonia s
Q8XV34 salmonella
Q8ZM1 salmonella
Q9HPQ2 halobacteri
Q48230 haemophilus
Q9H79 halobacteri
Q92N17 rhizobium m
Q9HPQ1 halobacteri
Q9F53 arabidopsis
Q9K7C5 bacillus ha
Q26707 trypanosoma
Q8SY10 drosophila
Q8ENV8 oceanobacil

| | | | | | | | | | | | | | |
|-----|---|-----|------|----|--------|--------------------|-----|---|-----|-----|----|--------|--------------------|
| 90 | 8 | 0.9 | 530 | 10 | Q8M041 | Q8W041 arabidopsis | 163 | 7 | 0.8 | 104 | 16 | Q8FD03 | Q8fd03 escherichia |
| 91 | 8 | 0.9 | 532 | 2 | Q2ZFG8 | Q2zfg8 azotobacter | 164 | 7 | 0.8 | 104 | 16 | Q8EK58 | Q8ek58 shewanella |
| 92 | 8 | 0.9 | 537 | 10 | Q9AYT3 | Q9ayt3 nicotiana t | 165 | 7 | 0.8 | 105 | 2 | Q9LBB1 | Q9lbb1 helicobacte |
| 93 | 8 | 0.9 | 539 | 16 | Q92XT9 | Q92xt9 rhizobium m | 166 | 7 | 0.8 | 105 | 2 | Q9RAL9 | Q9ral9 methylocyst |
| 94 | 8 | 0.9 | 553 | 2 | Q44493 | Q44493 azotobacter | 167 | 7 | 0.8 | 105 | 10 | Q8H3B5 | Q8h3b5 oryza sativ |
| 95 | 8 | 0.9 | 559 | 17 | Q9HRT5 | Q9hrt5 halobacteri | 168 | 7 | 0.8 | 105 | 16 | Q9KNZ5 | Q9knz5 vibrio chol |
| 96 | 8 | 0.9 | 559 | 16 | Q8YR01 | Q8yr01 anabaena sp | 169 | 7 | 0.8 | 107 | 12 | Q8S139 | Q8s139 human parvo |
| 97 | 8 | 0.9 | 589 | 16 | Q8YR01 | Q8yr01 anabaena sp | 169 | 7 | 0.8 | 107 | 12 | Q8S139 | Q8s139 human parvo |
| 98 | 8 | 0.9 | 607 | 11 | Q912Y4 | Q912y4 mus musculu | 170 | 7 | 0.8 | 107 | 16 | Q8NQ9 | Q8nq9 rhizobium l |
| 99 | 8 | 0.9 | 643 | 2 | Q68085 | Q68085 rhodobacter | 171 | 7 | 0.8 | 108 | 12 | Q8S143 | Q8s143 human parvo |
| 100 | 8 | 0.9 | 679 | 17 | Q8TYL7 | Q8tyl7 methanopyru | 172 | 7 | 0.8 | 109 | 13 | Q90WX7 | Q90wx7 rana catesb |
| 101 | 8 | 0.9 | 704 | 5 | Q9VIE2 | Q9vie2 desophila | 173 | 7 | 0.8 | 110 | 5 | Q20637 | Q20637 caenorhabdi |
| 102 | 8 | 0.9 | 802 | 16 | Q8CRU0 | Q8crj0 staphylococ | 174 | 7 | 0.8 | 110 | 16 | Q9KET3 | Q9ket3 bacillus ha |
| 103 | 8 | 0.9 | 825 | 2 | Q8KWA8 | Q8kwa8 ruegeria sp | 175 | 7 | 0.8 | 111 | 2 | Q06120 | Q06120 methylocyst |
| 104 | 8 | 0.9 | 850 | 2 | Q337V4 | Q337v4 bordetella | 176 | 7 | 0.8 | 112 | 10 | Q944K9 | Q944k9 arabidopsis |
| 105 | 8 | 0.9 | 850 | 2 | Q337V8 | Q337v8 bordetella | 177 | 7 | 0.8 | 112 | 16 | Q8TRE0 | Q8tre0 anabaena sp |
| 106 | 8 | 0.9 | 850 | 2 | Q337W0 | Q337w0 bordetella | 178 | 7 | 0.8 | 112 | 17 | Q27980 | Q27980 archaeoglob |
| 107 | 8 | 0.9 | 850 | 2 | Q337V9 | Q337v9 bordetella | 179 | 7 | 0.8 | 116 | 2 | Q9LBB7 | Q9lbb7 helicobacte |
| 108 | 8 | 0.9 | 850 | 2 | Q332V7 | Q332v7 bordetella | 180 | 7 | 0.8 | 117 | 2 | Q9LBA3 | Q9lba3 helicobacte |
| 109 | 8 | 0.9 | 850 | 2 | Q337W1 | Q337w1 bordetella | 181 | 7 | 0.8 | 120 | 9 | Q8W5Z8 | Q8w5z8 bacterioph |
| 110 | 8 | 0.9 | 891 | 2 | Q8RMZ9 | Q8rmz9 azotobacter | 182 | 7 | 0.8 | 121 | 2 | Q9L362 | Q9l362 helicobacte |
| 111 | 8 | 0.9 | 993 | 16 | Q8YQ05 | Q8yyq5 anabaena sp | 183 | 7 | 0.8 | 121 | 2 | Q9L362 | Q9l362 helicobacte |
| 112 | 8 | 0.9 | 997 | 2 | Q44492 | Q44492 azotobacter | 184 | 7 | 0.8 | 123 | 6 | Q62695 | Q62695 oryctolagus |
| 113 | 8 | 0.9 | 997 | 2 | Q44495 | Q44495 azotobacter | 185 | 7 | 0.8 | 123 | 11 | Q8CFT8 | Q8cft8 mus musculu |
| 114 | 8 | 0.9 | 1000 | 15 | Q8JBT8 | Q8jbt8 human immun | 186 | 7 | 0.8 | 126 | 2 | Q9LBB9 | Q9lbb9 helicobacte |
| 115 | 8 | 0.9 | 1003 | 15 | Q8AK15 | Q8aki5 human immun | 187 | 7 | 0.8 | 126 | 17 | Q9HRR7 | Q9hrr7 halobacteri |
| 116 | 8 | 0.9 | 1055 | 2 | Q43892 | Q43892 actinobacil | 188 | 7 | 0.8 | 127 | 16 | Q8G4L1 | Q8g4l1 bifidobacte |
| 117 | 8 | 0.9 | 1112 | 16 | Q92UV3 | Q92uv3 rhizobium m | 189 | 7 | 0.8 | 128 | 5 | Q8MW17 | Q8mw17 boltenia vi |
| 118 | 8 | 0.9 | 1168 | 2 | Q9AGD2 | Q9agd2 azotobacter | 190 | 7 | 0.8 | 130 | 2 | Q9LBS2 | Q9lbs2 helicobacte |
| 119 | 8 | 0.9 | 1208 | 16 | Q9PFI9 | Q9pfi9 xylella fas | 191 | 7 | 0.8 | 130 | 16 | Q8F0G1 | Q8f0g1 leptospira |
| 120 | 8 | 0.9 | 1296 | 16 | Q9P9W1 | Q9p9w1 xylella fas | 192 | 7 | 0.8 | 132 | 2 | Q9LBB9 | Q9lbb9 helicobacte |
| 121 | 8 | 0.9 | 1417 | 16 | Q8YTO9 | Q8yto9 anabaena sp | 193 | 7 | 0.8 | 133 | 10 | Q04338 | Q04338 arabidopsis |
| 122 | 8 | 0.9 | 1706 | 2 | Q8L469 | Q8l469 bordetella | 194 | 7 | 0.8 | 133 | 10 | Q39194 | Q39194 arabidopsis |
| 123 | 8 | 0.9 | 1736 | 10 | Q23025 | Q23025 arabidopsis | 195 | 7 | 0.8 | 133 | 16 | Q97E52 | Q97e52 clostridium |
| 124 | 8 | 0.9 | 1741 | 16 | F73817 | F73817 synochocyst | 196 | 7 | 0.8 | 134 | 12 | Q8S178 | Q8s178 human parvo |
| 125 | 8 | 0.9 | 1839 | 2 | Q44496 | Q44496 azotobacter | 197 | 7 | 0.8 | 135 | 2 | Q9LBB5 | Q9lbb5 helicobacte |
| 126 | 8 | 0.9 | 1990 | 16 | Q8U7N7 | Q8u7n7 agrobacteri | 198 | 7 | 0.8 | 135 | 2 | Q9LBB7 | Q9lbb7 helicobacte |
| 127 | 8 | 0.9 | 2361 | 5 | Q8IEJ4 | Q8iej4 plasmodium | 199 | 7 | 0.8 | 136 | 2 | Q9K4W1 | Q9k4w1 streptomyce |
| 128 | 8 | 0.9 | 2747 | 2 | Q2L800 | Q2l800 aeromonas s | 200 | 7 | 0.8 | 138 | 11 | Q8BW15 | Q8bw15 mus musculu |
| 129 | 8 | 0.9 | 3083 | 16 | Q8YL10 | Q8yll0 anabaena sp | 201 | 7 | 0.8 | 139 | 16 | Q8ENQ5 | Q8enq5 oceanobacil |
| 130 | 7 | 0.8 | 17 | 6 | Q28459 | Q28459 mirounga an | 202 | 7 | 0.8 | 140 | 2 | Q9LBB9 | Q9lbb9 helicobacte |
| 131 | 7 | 0.8 | 17 | 6 | Q27943 | Q27943 arctocephal | 203 | 7 | 0.8 | 140 | 2 | Q9LBB9 | Q9lbb9 helicobacte |
| 132 | 7 | 0.8 | 17 | 6 | Q28246 | Q28246 canis famil | 204 | 7 | 0.8 | 140 | 2 | Q9LBB7 | Q9lbb7 helicobacte |
| 133 | 7 | 0.8 | 17 | 6 | Q28455 | Q28455 leptonyctot | 205 | 7 | 0.8 | 140 | 12 | Q9J891 | Q9j891 spodoptera |
| 134 | 7 | 0.8 | 17 | 6 | Q28817 | Q28817 pioca vitul | 206 | 7 | 0.8 | 142 | 4 | Q8WYQ3 | Q8wyq3 homo sapien |
| 135 | 7 | 0.8 | 17 | 6 | Q28445 | Q28445 hydruaga le | 207 | 7 | 0.8 | 143 | 2 | Q9LBB7 | Q9lbb7 helicobacte |
| 136 | 7 | 0.8 | 31 | 6 | Q28493 | Q28493 mirounga le | 208 | 7 | 0.8 | 145 | 2 | Q9LBB9 | Q9lbb9 helicobacte |
| 137 | 7 | 0.8 | 63 | 5 | Q9GSH2 | Q9gsh2 bombyx mori | 209 | 7 | 0.8 | 147 | 10 | Q9C674 | Q9c674 arabidopsis |
| 138 | 7 | 0.8 | 63 | 16 | Q92JF1 | Q92jfi rickettsia | 210 | 7 | 0.8 | 147 | 16 | Q53898 | Q53898 streptomyce |
| 139 | 7 | 0.8 | 73 | 9 | Q8SDW7 | Q8sdw7 bacterioph | 211 | 7 | 0.8 | 149 | 5 | Q19502 | Q19502 caenorhabdi |
| 140 | 7 | 0.8 | 73 | 9 | Q9G032 | Q9g032 bacterioph | 212 | 7 | 0.8 | 151 | 2 | Q9LBB5 | Q9lbb5 helicobacte |
| 141 | 7 | 0.8 | 73 | 9 | Q9MBS7 | Q9mbs7 staphylococ | 213 | 7 | 0.8 | 151 | 4 | Q9Y6H1 | Q9y6h1 homo sapien |
| 142 | 7 | 0.8 | 73 | 16 | Q932B3 | Q932b3 staphylococ | 214 | 7 | 0.8 | 151 | 17 | Q97Z82 | Q97z82 sulfolobus |
| 143 | 7 | 0.8 | 75 | 8 | Q9G9R8 | Q9gr8 thunnus ton | 215 | 7 | 0.8 | 153 | 11 | Q9D1L0 | Q9d1l0 mus musculu |
| 144 | 7 | 0.8 | 75 | 8 | Q9G101 | Q9g101 thunnus obe | 216 | 7 | 0.8 | 157 | 16 | Q8FJZ0 | Q8fjz0 xanthomonas |
| 145 | 7 | 0.8 | 75 | 8 | Q9G6D1 | Q9g6d1 thunnus ala | 217 | 7 | 0.8 | 161 | 11 | Q8K3S8 | Q8k3s8 mus musculu |
| 146 | 7 | 0.8 | 75 | 8 | Q9G6C5 | Q9g6c5 thunnus alb | 218 | 7 | 0.8 | 164 | 10 | Q9P57 | Q9p57 oryza sativ |
| 147 | 7 | 0.8 | 75 | 8 | Q9G6C7 | Q9g6c7 thunnus mac | 219 | 7 | 0.8 | 167 | 4 | Q8TEP2 | Q8tep2 homo sapien |
| 148 | 7 | 0.8 | 75 | 8 | Q9G6C9 | Q9g6c9 thunnus thy | 220 | 7 | 0.8 | 168 | 2 | Q59609 | Q59609 neisseria g |
| 149 | 7 | 0.8 | 75 | 8 | Q9G9S0 | Q9g9s0 thunnus atl | 221 | 7 | 0.8 | 169 | 12 | Q912Q3 | Q912q3 human parvo |
| 150 | 7 | 0.8 | 75 | 8 | Q9G6D3 | Q9g6d3 thunnus thy | 222 | 7 | 0.8 | 169 | 12 | Q912C4 | Q912c4 human eryth |
| 151 | 7 | 0.8 | 79 | 16 | Q9A3P6 | Q9a3p6 caulobacter | 223 | 7 | 0.8 | 169 | 12 | Q912P1 | Q912p1 human parvo |
| 152 | 7 | 0.8 | 81 | 16 | Q8U5J3 | Q8u5j3 agrobacteri | 224 | 7 | 0.8 | 169 | 12 | Q912Q1 | Q912q1 human parvo |
| 153 | 7 | 0.8 | 85 | 16 | Q9A4J2 | Q9a4j2 caulobacter | 225 | 7 | 0.8 | 169 | 12 | Q912N7 | Q912n7 human parvo |
| 154 | 7 | 0.8 | 90 | 11 | Q8R0H5 | Q8r0h5 pseudomonas | 226 | 7 | 0.8 | 169 | 12 | Q912P9 | Q912p9 human parvo |
| 155 | 7 | 0.8 | 92 | 10 | Q93WB4 | Q93wb4 arabidopsis | 227 | 7 | 0.8 | 169 | 12 | Q912C1 | Q912c1 human eryth |
| 156 | 7 | 0.8 | 98 | 10 | Q8W233 | Q8w233 zea mays (m | 228 | 7 | 0.8 | 169 | 12 | Q912P6 | Q912p6 human parvo |
| 157 | 7 | 0.8 | 102 | 12 | Q8S128 | Q8s128 human parvo | 229 | 7 | 0.8 | 169 | 12 | Q912N9 | Q912n9 human parvo |
| 158 | 7 | 0.8 | 102 | 16 | Q8PL98 | Q8pl98 xanthomonas | 230 | 7 | 0.8 | 169 | 12 | Q912Q5 | Q912q5 human parvo |
| 159 | 7 | 0.8 | 102 | 16 | Q8P9H7 | Q8p9h7 xanthomonas | 231 | 7 | 0.8 | 169 | 12 | Q912N2 | Q912n2 human parvo |
| 160 | 7 | 0.8 | 104 | 12 | Q8S152 | Q8s152 human parvo | 232 | 7 | 0.8 | 169 | 12 | Q912D0 | Q912d0 human eryth |
| 161 | 7 | 0.8 | 104 | 16 | Q8Z1X8 | Q8z1x8 salmonella | 233 | 7 | 0.8 | 169 | 12 | Q912C6 | Q912c6 human eryth |
| 162 | 7 | 0.8 | 104 | 16 | Q8ZJAI | Q8zjai yersinia pe | 234 | 7 | 0.8 | 169 | 12 | Q912N5 | Q912n5 human parvo |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|----|--------|--------------------|-----|---|-----|-----|----|---------|---------------------|
| 236 | 7 | 0.8 | 169 | 12 | Q912P4 | Q912P4 human parvo | 309 | 7 | 0.8 | 249 | 16 | Q8U6D0 | Q8U6D0 agrobacteri |
| 237 | 7 | 0.8 | 170 | 5 | Q9NER7 | Q9ner7 caenorhabdi | 310 | 7 | 0.8 | 249 | 16 | Q8RKH4 | Q8rkH4 streptomyce |
| 238 | 7 | 0.8 | 172 | 16 | Q9ZV01 | Q9zV01 rhizobium m | 311 | 7 | 0.8 | 250 | 12 | Q8DKM8 | Q8dkm8 human parvo |
| 239 | 7 | 0.8 | 175 | 2 | Q8GQV7 | Q8gQV7 thibacillu | 312 | 7 | 0.8 | 250 | 12 | Q9DPM5 | Q9dpm5 human parvo |
| 240 | 7 | 0.8 | 175 | 2 | Q8GQY3 | Q8gQY3 thibacillu | 313 | 7 | 0.8 | 250 | 12 | Q9DPM5 | Q9dpm5 human parvo |
| 241 | 7 | 0.8 | 175 | 16 | Q8S6X0 | Q8s6X0 streptococ | 314 | 7 | 0.8 | 250 | 12 | Q9DPM9 | Q9dpm9 human parvo |
| 242 | 7 | 0.8 | 175 | 16 | Q8S1F2 | Q8s1F2 streptococ | 315 | 7 | 0.8 | 250 | 12 | Q9DPP1 | Q9dpp1 human parvo |
| 243 | 7 | 0.8 | 176 | 12 | Q8S182 | Q8s182 human parvo | 316 | 7 | 0.8 | 250 | 12 | Q9DPP2 | Q9dpp2 human parvo |
| 244 | 7 | 0.8 | 177 | 10 | Q8W1H0 | Q8w1H0 triticum ae | 317 | 7 | 0.8 | 250 | 12 | Q9DPP3 | Q9dpp3 human parvo |
| 245 | 7 | 0.8 | 177 | 16 | Q8DFN3 | Q8dfN3 vibrio vuln | 318 | 7 | 0.8 | 250 | 12 | Q9DPP8 | Q9dpp8 human parvo |
| 246 | 7 | 0.8 | 178 | 2 | Q8GQY4 | Q8gQY4 thibacillu | 319 | 7 | 0.8 | 250 | 12 | Q9DPP1 | Q9dpp1 human parvo |
| 247 | 7 | 0.8 | 178 | 16 | Q9F3M4 | Q9f3M4 streptomyce | 320 | 7 | 0.8 | 250 | 12 | Q9DPP9 | Q9dpp9 human parvo |
| 248 | 7 | 0.8 | 179 | 12 | Q8S1Y2 | Q8s1Y2 human parvo | 321 | 7 | 0.8 | 250 | 12 | Q9DPP4 | Q9dpp4 human parvo |
| 249 | 7 | 0.8 | 181 | 2 | Q68403 | Q68403 escherichia | 322 | 7 | 0.8 | 250 | 12 | Q9DPP7 | Q9dpp7 human parvo |
| 250 | 7 | 0.8 | 181 | 2 | Q70070 | Q70070 escherichia | 323 | 7 | 0.8 | 250 | 12 | Q9DPP4 | Q9dpp4 human parvo |
| 251 | 7 | 0.8 | 181 | 2 | Q68404 | Q68404 escherichia | 324 | 7 | 0.8 | 250 | 12 | Q9DPP2 | Q9dpp2 human parvo |
| 252 | 7 | 0.8 | 181 | 8 | Q9TLU6 | Q9tlu6 cyanidium c | 325 | 7 | 0.8 | 250 | 12 | Q9DPP6 | Q9dpp6 human parvo |
| 253 | 7 | 0.8 | 181 | 12 | Q8S120 | Q8s120 human parvo | 326 | 7 | 0.8 | 250 | 12 | Q9DPP0 | Q9dpp0 human parvo |
| 254 | 7 | 0.8 | 182 | 12 | Q8S147 | Q8s147 human parvo | 327 | 7 | 0.8 | 250 | 12 | Q9DPP7 | Q9dpp7 human parvo |
| 255 | 7 | 0.8 | 186 | 12 | Q8S163 | Q8s163 human parvo | 328 | 7 | 0.8 | 250 | 12 | Q9DPP6 | Q9dpp6 human parvo |
| 256 | 7 | 0.8 | 186 | 12 | Q8S132 | Q8s132 human parvo | 329 | 7 | 0.8 | 250 | 12 | Q9DPP3 | Q9dpp3 human parvo |
| 257 | 7 | 0.8 | 187 | 16 | Q8XNW0 | Q8xNW0 clostridium | 330 | 7 | 0.8 | 250 | 12 | Q9DPP0 | Q9dpp0 human parvo |
| 258 | 7 | 0.8 | 187 | 16 | Q8D5Y0 | Q8d5Y0 vibrio vuln | 331 | 7 | 0.8 | 251 | 16 | Q8ZRM2 | Q8zrm2 salmonella |
| 259 | 7 | 0.8 | 189 | 12 | Q8JLD1 | Q8jld1 ectromelia | 332 | 7 | 0.8 | 251 | 16 | Q8Z983 | Q8z983 salmonella |
| 260 | 7 | 0.8 | 189 | 16 | Q8ZFF8 | Q8zff8 versinia pe | 333 | 7 | 0.8 | 253 | 16 | Q8NVX1 | Q8nvX1 staphylococ |
| 261 | 7 | 0.8 | 190 | 17 | Q8ZX11 | Q8zx11 pyrobaculum | 334 | 7 | 0.8 | 254 | 16 | Q8XEE8 | Q8xee8 escherichia |
| 262 | 7 | 0.8 | 192 | 16 | Q8XKV0 | Q8xKV0 clostridium | 335 | 7 | 0.8 | 256 | 16 | Q8XP43 | Q8xp43 clostridium |
| 263 | 7 | 0.8 | 198 | 12 | Q8S193 | Q8s193 human parvo | 336 | 7 | 0.8 | 259 | 10 | Q8R730 | Q8r730 arum macula |
| 264 | 7 | 0.8 | 199 | 10 | Q8VXD6 | Q8vxd6 nanochloro | 337 | 7 | 0.8 | 260 | 2 | Q8RQU3 | Q8rqu3 thermus the |
| 265 | 7 | 0.8 | 199 | 16 | P73231 | P73231 synechocyst | 338 | 7 | 0.8 | 260 | 5 | Q9S5W12 | Q9s5w12 anopheles a |
| 266 | 7 | 0.8 | 202 | 5 | Q8LJZ2 | Q8ljz2 plasmodium | 339 | 7 | 0.8 | 260 | 10 | Q38731 | Q38731 arum macula |
| 267 | 7 | 0.8 | 204 | 2 | Q9Z3G4 | Q9z3G4 synechococ | 340 | 7 | 0.8 | 261 | 8 | Q94SD6 | Q94sd6 dactylopten |
| 268 | 7 | 0.8 | 211 | 10 | Q9LQB6 | Q9lqb6 arabidopsis | 341 | 7 | 0.8 | 261 | 8 | Q94SE9 | Q94se9 mastacemba |
| 269 | 7 | 0.8 | 211 | 12 | Q8QY8 | Q8qy8 foot-and-mo | 342 | 7 | 0.8 | 261 | 8 | Q94T14 | Q94t14 caelolinchu |
| 270 | 7 | 0.8 | 213 | 16 | Q9CBJ1 | Q9cbj1 mycobacteri | 343 | 7 | 0.8 | 261 | 8 | Q94T14 | Q94t14 iijimaia dof |
| 271 | 7 | 0.8 | 215 | 17 | Q8Z2Y1 | Q8z2Y1 pyrobaculum | 344 | 7 | 0.8 | 261 | 8 | Q94S80 | Q94s80 pagrus majo |
| 272 | 7 | 0.8 | 216 | 16 | Q8RUL4 | Q8rUL4 deinococcus | 345 | 7 | 0.8 | 261 | 8 | Q94S18 | Q94s18 antagonia c |
| 273 | 7 | 0.8 | 217 | 16 | Q8U8S1 | Q8u8S1 agrobacteri | 346 | 7 | 0.8 | 261 | 8 | Q94XJ1 | Q94xj1 neoceratodu |
| 274 | 7 | 0.8 | 218 | 16 | Q9P6F5 | Q9p6F5 xyella fas | 347 | 7 | 0.8 | 261 | 8 | Q94TB3 | Q94tb3 neoscopolus |
| 275 | 7 | 0.8 | 219 | 12 | Q9W9Q6 | Q9w9Q6 human parvo | 348 | 7 | 0.8 | 261 | 8 | Q94SR6 | Q94sr6 rondefletia |
| 276 | 7 | 0.8 | 219 | 12 | Q9W9B9 | Q9w9B9 human parvo | 349 | 7 | 0.8 | 261 | 8 | Q94T19 | Q94t19 chauiodous |
| 277 | 7 | 0.8 | 219 | 12 | Q9WAS9 | Q9was9 human parvo | 350 | 7 | 0.8 | 261 | 8 | Q94T87 | Q94t87 diaphus spl |
| 278 | 7 | 0.8 | 219 | 16 | Q9HWX3 | Q9hwX3 pseudomonas | 351 | 7 | 0.8 | 261 | 8 | Q94SQ3 | Q94sq3 hoplostethu |
| 279 | 7 | 0.8 | 219 | 16 | P96437 | P96437 rhizobium m | 352 | 7 | 0.8 | 261 | 8 | Q9G6P4 | Q9g6P4 polymixia j |
| 280 | 7 | 0.8 | 219 | 16 | Q66733 | Q66733 aquifex aco | 353 | 7 | 0.8 | 261 | 8 | Q8SF50 | Q8sf50 ictalurus p |
| 281 | 7 | 0.8 | 220 | 17 | Q26974 | Q26974 methanobact | 354 | 7 | 0.8 | 261 | 8 | Q94TF0 | Q94tf0 chlorophtha |
| 282 | 7 | 0.8 | 221 | 2 | Q9ZBA7 | Q9zba7 streptomyce | 355 | 7 | 0.8 | 261 | 8 | Q9B603 | Q9b603 plecoglossu |
| 283 | 7 | 0.8 | 223 | 5 | Q95TD5 | Q95td5 drosophila | 356 | 7 | 0.8 | 261 | 8 | Q94SW8 | Q94sw8 exocoetus v |
| 284 | 7 | 0.8 | 225 | 16 | Q8PD09 | Q8pd09 xanthomonas | 357 | 7 | 0.8 | 261 | 8 | Q94T48 | Q94t48 zu cristatu |
| 285 | 7 | 0.8 | 227 | 9 | Q8HA71 | Q8ha71 bacterioph | 358 | 7 | 0.8 | 261 | 8 | Q94SK1 | Q94sk1 zenopsis ne |
| 286 | 7 | 0.8 | 227 | 16 | Q9HW21 | Q9hw21 pseudomonas | 359 | 7 | 0.8 | 261 | 8 | Q9B629 | Q9b629 anguilla ja |
| 287 | 7 | 0.8 | 228 | 16 | Q8PGF3 | Q8pgf3 xanthomonas | 360 | 7 | 0.8 | 261 | 8 | Q94SS9 | Q94ss9 danactetic |
| 288 | 7 | 0.8 | 229 | 9 | Q9ZX28 | Q9zx28 mycobacteri | 361 | 7 | 0.8 | 261 | 8 | Q94T38 | Q94t38 polymixia l |
| 289 | 7 | 0.8 | 231 | 12 | Q9WAU4 | Q9wau4 human parvo | 362 | 7 | 0.8 | 261 | 8 | Q94TC6 | Q94tc6 sauridia und |
| 290 | 7 | 0.8 | 231 | 12 | Q9WAT0 | Q9wat0 human parvo | 363 | 7 | 0.8 | 261 | 8 | Q9MIA3 | Q9mia3 typhlonecte |
| 291 | 7 | 0.8 | 231 | 12 | Q9W8V3 | Q9w8V3 human parvo | 364 | 7 | 0.8 | 261 | 8 | Q94SL4 | Q94sl4 zeus faber |
| 292 | 7 | 0.8 | 231 | 12 | Q9W929 | Q9w929 human parvo | 365 | 7 | 0.8 | 261 | 8 | Q9BSU9 | Q9bsu9 aulopus jap |
| 293 | 7 | 0.8 | 231 | 12 | Q9WAS8 | Q9was8 human parvo | 366 | 7 | 0.8 | 261 | 8 | Q94SY1 | Q94sy1 colorablis s |
| 294 | 7 | 0.8 | 231 | 12 | Q9WAS5 | Q9was5 human parvo | 367 | 7 | 0.8 | 261 | 8 | Q94T61 | Q94t61 trachipteru |
| 295 | 7 | 0.8 | 233 | 16 | Q9EWM6 | Q9ewm6 streptomyce | 368 | 7 | 0.8 | 261 | 8 | Q94SM7 | Q94sm7 myripristis |
| 296 | 7 | 0.8 | 233 | 10 | Q9FFH2 | Q9ffh2 arabidopsis | 369 | 7 | 0.8 | 261 | 8 | Q94TD8 | Q94td8 harpadon mi |
| 297 | 7 | 0.8 | 237 | 2 | Q33630 | Q33630 saccharopol | 370 | 7 | 0.8 | 261 | 8 | Q8HQP0 | Q8hqp0 acipenser t |
| 298 | 7 | 0.8 | 238 | 2 | Q69260 | Q69260 borrelia re | 371 | 7 | 0.8 | 261 | 8 | Q8HQ15 | Q8hq15 amia calva |
| 299 | 7 | 0.8 | 241 | 16 | Q9RZM1 | Q9rzm1 deinococcus | 372 | 7 | 0.8 | 261 | 8 | Q8HMU9 | Q8hmu9 gymnothorax |
| 300 | 7 | 0.8 | 242 | 2 | Q9EX91 | Q9ex91 petrotoga m | 373 | 7 | 0.8 | 261 | 8 | Q8HMT7 | Q8hmt7 erpetoichth |
| 301 | 7 | 0.8 | 242 | 5 | Q8T925 | Q8t925 tetrahymena | 374 | 7 | 0.8 | 261 | 8 | Q8HMS7 | Q8hms7 polypeterus |
| 302 | 7 | 0.8 | 242 | 16 | Q8ZKK9 | Q8zkk9 salmonella | 375 | 7 | 0.8 | 261 | 8 | Q8HMR7 | Q8hmr7 polypeterus |
| 303 | 7 | 0.8 | 243 | 3 | Q96UE4 | Q96ue4 neurospora | 376 | 7 | 0.8 | 261 | 8 | Q8HMQ4 | Q8hmQ4 polyodon sp |
| 304 | 7 | 0.8 | 243 | 10 | Q9CAX0 | Q9cax0 arabidopsis | 377 | 7 | 0.8 | 261 | 8 | Q8HMP1 | Q8hmp1 scaphirhinc |
| 305 | 7 | 0.8 | 244 | 10 | Q949S4 | Q949s4 arabidopsis | 378 | 7 | 0.8 | 261 | 8 | Q8HMH6 | Q8hmh6 basozetous |
| 306 | 7 | 0.8 | 244 | 10 | Q8L8V6 | Q8l8v6 arabidopsis | 379 | 7 | 0.8 | 261 | 8 | Q8HMH6 | Q8hmh6 lota lota. |
| 307 | 7 | 0.8 | 246 | 8 | Q95PD3 | Q95fd3 riccia flui | 380 | 7 | 0.8 | 261 | 8 | Q8HM73 | Q8hm73 lophius lit |
| 308 | 7 | 0.8 | 248 | 2 | Q87368 | Q87368 staphylococ | 381 | 7 | 0.8 | 261 | 8 | Q8HM60 | Q8hm60 lophius ame |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|----|--------|---------------------|-----|---|-----|-----|----|--------|---------------------|
| 382 | 7 | 0.8 | 261 | 8 | Q8HM21 | Q8hm21 caulophryne | 455 | 7 | 0.8 | 311 | 16 | Q9RKA0 | Q9rka0 streptomyc |
| 383 | 7 | 0.8 | 261 | 8 | Q8HM08 | Q8hm08 melanocetus | 456 | 7 | 0.8 | 311 | 17 | Q9VDE9 | Q9vde9 aeropyrum p |
| 384 | 7 | 0.8 | 261 | 8 | Q8HLZ5 | Q8hlz5 melanotaeni | 457 | 7 | 0.8 | 312 | 4 | Q92977 | Q92977 homo sapien |
| 385 | 7 | 0.8 | 261 | 8 | Q8HLW9 | Q8hlw9 yersias lat | 458 | 7 | 0.8 | 312 | 4 | Q99620 | Q99620 homo sapien |
| 386 | 7 | 0.8 | 261 | 8 | Q8HLV7 | Q8hlv7 gambusia af | 459 | 7 | 0.8 | 312 | 5 | Q19261 | Q19261 caenorhabdi |
| 387 | 7 | 0.8 | 261 | 8 | Q8HLV4 | Q8hlv4 testotoma r | 460 | 7 | 0.8 | 314 | 2 | Q52240 | Q52240 staphylococ |
| 388 | 7 | 0.8 | 261 | 8 | Q8HLT1 | Q8hlt1 eutaeniopho | 461 | 7 | 0.8 | 314 | 10 | Q9M9A3 | Q9m9a3 arabidopsis |
| 389 | 7 | 0.8 | 261 | 8 | Q8HLO5 | Q8hlo5 diretnoides | 462 | 7 | 0.8 | 314 | 16 | Q8XNN5 | Q8xnn5 clostridium |
| 390 | 7 | 0.8 | 261 | 8 | Q8HLP2 | Q8hlp2 directus ar | 463 | 7 | 0.8 | 315 | 9 | Q8SDT6 | Q8sd6 bacterioph |
| 391 | 7 | 0.8 | 261 | 8 | Q8HLM9 | Q8hlm9 anomalops k | 464 | 7 | 0.8 | 315 | 16 | Q931X0 | Q931x0 staphylococ |
| 392 | 7 | 0.8 | 261 | 8 | Q8HLL6 | Q8hll6 monocentris | 465 | 7 | 0.8 | 315 | 16 | Q8FLW0 | Q8flw0 corynebacte |
| 393 | 7 | 0.8 | 261 | 8 | Q8HLK3 | Q8hlk3 beryx decad | 466 | 7 | 0.8 | 315 | 16 | Q8ECM6 | Q8ecm6 shewanella |
| 394 | 7 | 0.8 | 261 | 8 | Q8HLJ0 | Q8hlj0 ostichthys | 467 | 7 | 0.8 | 316 | 10 | Q941A9 | Q941a9 arabidopsis |
| 395 | 7 | 0.8 | 261 | 8 | Q8HLH7 | Q8hlh7 sargocentro | 468 | 7 | 0.8 | 317 | 16 | Q9KEH0 | Q9keh0 bacillus ha |
| 396 | 7 | 0.8 | 261 | 8 | Q8HLG4 | Q8hlg4 parazen pac | 469 | 7 | 0.8 | 318 | 10 | Q64946 | Q64946 cumunis sat |
| 397 | 7 | 0.8 | 261 | 8 | Q8HLD8 | Q8hld8 allocyttus r | 470 | 7 | 0.8 | 319 | 16 | Q912F8 | Q912f8 pseudomonas |
| 398 | 7 | 0.8 | 261 | 8 | Q8HLC5 | Q8hlc5 neocytus r | 471 | 7 | 0.8 | 319 | 16 | Q8YU40 | Q8yu40 anabaena sp |
| 399 | 7 | 0.8 | 261 | 8 | Q8HKZ6 | Q8hkz6 emelichthy | 472 | 7 | 0.8 | 320 | 10 | Q9CAA2 | Q9caa2 arabidopsis |
| 400 | 7 | 0.8 | 261 | 8 | Q8HKY3 | Q8hky3 pterocaeio | 473 | 7 | 0.8 | 320 | 10 | Q94CK9 | Q94ck9 arabidopsis |
| 401 | 7 | 0.8 | 261 | 8 | Q8HKT1 | Q8hkt1 salarias fa | 474 | 7 | 0.8 | 320 | 16 | Q8YNL3 | Q8ynl3 anabaena sp |
| 402 | 7 | 0.8 | 261 | 8 | Q8HKQ5 | Q8hkq5 aspaema min | 475 | 7 | 0.8 | 321 | 16 | Q8DLI7 | Q8dl17 synechococ |
| 403 | 7 | 0.8 | 261 | 8 | Q8HKP2 | Q8hkp2 rhyacichthy | 476 | 7 | 0.8 | 324 | 16 | Q910D3 | Q910d3 pseudomonas |
| 404 | 7 | 0.8 | 261 | 8 | Q8HKM9 | Q8hkm9 rheotris ac | 477 | 7 | 0.8 | 324 | 16 | Q92PN4 | Q92pn4 rhizobium m |
| 405 | 7 | 0.8 | 261 | 8 | Q8HKL6 | Q8hkl6 sufflamen f | 478 | 7 | 0.8 | 325 | 5 | Q8T896 | Q8t896 ciona intes |
| 406 | 7 | 0.8 | 261 | 8 | Q8HCV9 | Q8hcv9 fugu rubrip | 479 | 7 | 0.8 | 325 | 17 | Q8TV21 | Q8tv21 methanopyru |
| 407 | 7 | 0.8 | 261 | 16 | Q8EQZ4 | Q8eqz4 oceanobacil | 480 | 7 | 0.8 | 325 | 17 | Q8TVV2 | Q8tvv2 methanopyru |
| 408 | 7 | 0.8 | 262 | 12 | Q8B9H8 | Q8b9h8 rachiplusia | 481 | 7 | 0.8 | 326 | 3 | Q02461 | Q02461 saccharomyc |
| 409 | 7 | 0.8 | 262 | 16 | Q69841 | Q69841 streptomyc | 482 | 7 | 0.8 | 328 | 5 | Q8SZR6 | Q8szr6 drosophila |
| 410 | 7 | 0.8 | 264 | 13 | Q9W7F6 | Q9w7f6 polyodon sp | 483 | 7 | 0.8 | 329 | 2 | P72466 | P72466 streptomyc |
| 411 | 7 | 0.8 | 267 | 10 | Q8RYD3 | Q8ryd3 arabidopsis | 484 | 7 | 0.8 | 331 | 5 | Q9VJ31 | Q9vj31 drosophila |
| 412 | 7 | 0.8 | 269 | 10 | Q9SMC6 | Q9smc6 capicum an | 485 | 7 | 0.8 | 331 | 11 | Q8R4T3 | Q8rt43 rattus norv |
| 413 | 7 | 0.8 | 269 | 16 | P95001 | P95001 mycobacteri | 486 | 7 | 0.8 | 334 | 16 | Q9RL53 | Q9rl53 streptomyc |
| 414 | 7 | 0.8 | 270 | 10 | P9XG37 | P9xg37 plisum sativ | 487 | 7 | 0.8 | 335 | 2 | Q9L6C3 | Q9l6c3 streptomyc |
| 415 | 7 | 0.8 | 272 | 17 | Q8U3F7 | Q8u3f7 pyrococcus | 488 | 7 | 0.8 | 335 | 5 | Q9RKZ8 | Q9rkz8 streptomyc |
| 416 | 7 | 0.8 | 274 | 2 | Q9XS58 | Q9xs58 streptomyc | 489 | 7 | 0.8 | 336 | 5 | Q9VQL9 | Q9vql9 drosophila |
| 417 | 7 | 0.8 | 274 | 2 | Q9S3V2 | Q9s3v2 zymomonas m | 490 | 7 | 0.8 | 337 | 16 | Q8P4X0 | Q8p4x0 xanthomonas |
| 418 | 7 | 0.8 | 276 | 10 | Q8W4T7 | Q8w4t7 medicago tr | 491 | 7 | 0.8 | 338 | 16 | Q9KQ11 | Q9kq11 vibrio chol |
| 419 | 7 | 0.8 | 276 | 10 | Q40750 | Q40750 phaseolus a | 492 | 7 | 0.8 | 338 | 16 | Q9A6I1 | Q9a6i1 caulobacter |
| 420 | 7 | 0.8 | 278 | 16 | Q9I3H1 | Q9i3hi pseudomonas | 493 | 7 | 0.8 | 338 | 17 | Q58526 | Q58526 pyrococcus |
| 421 | 7 | 0.8 | 279 | 5 | Q8MR71 | Q8mr71 drosophila | 494 | 7 | 0.8 | 339 | 16 | Q987Z3 | Q987z3 rhizobium l |
| 422 | 7 | 0.8 | 279 | 16 | Q9A873 | Q9a873 caulobacter | 495 | 7 | 0.8 | 339 | 16 | Q8YGC6 | Q8ygc6 bruceella me |
| 423 | 7 | 0.8 | 280 | 16 | Q8E460 | Q8e460 streptococ | 496 | 7 | 0.8 | 340 | 11 | Q35125 | Q35125 mus musculu |
| 424 | 7 | 0.8 | 280 | 16 | Q8DYK4 | Q8dyk4 streptococ | 497 | 7 | 0.8 | 342 | 2 | Q8KZB7 | Q8kzb7 helicobacte |
| 425 | 7 | 0.8 | 281 | 16 | Q8Y4S3 | Q8y4s3 listeria mo | 498 | 7 | 0.8 | 343 | 4 | Q8NGK6 | Q8ngk6 homo sapien |
| 426 | 7 | 0.8 | 282 | 5 | Q18092 | Q18092 caenorhabdi | 499 | 7 | 0.8 | 343 | 5 | Q18636 | Q18636 musca domes |
| 427 | 7 | 0.8 | 284 | 16 | Q8XZQ4 | Q8xzq4 ralstonia s | 500 | 7 | 0.8 | 345 | 16 | Q98KQ2 | Q98kq2 rhizobium l |
| 428 | 7 | 0.8 | 284 | 16 | Q8DLF6 | Q8dlf6 synechococ | 501 | 7 | 0.8 | 346 | 8 | Q9ML74 | Q9ml74 bonasa sewe |
| 429 | 7 | 0.8 | 287 | 2 | Q8GPH2 | Q8gph2 pantoea agg | 502 | 7 | 0.8 | 346 | 16 | Q9ML73 | Q9ml73 bonasa umbe |
| 430 | 7 | 0.8 | 287 | 16 | Q8ELP0 | Q8elp0 oceanobacil | 503 | 7 | 0.8 | 346 | 16 | Q9RKT4 | Q9rkt4 streptomyc |
| 431 | 7 | 0.8 | 290 | 2 | Q52218 | Q52218 klebsiella | 504 | 7 | 0.8 | 348 | 16 | Q9K9N0 | Q9k9n0 bacillus ha |
| 432 | 7 | 0.8 | 290 | 16 | Q8D4G3 | Q8d4g3 vibrio vuln | 505 | 7 | 0.8 | 349 | 5 | Q22512 | Q22512 caenorhabdi |
| 433 | 7 | 0.8 | 291 | 16 | Q8ZKC9 | Q8zkc9 salmonella | 506 | 7 | 0.8 | 349 | 16 | Q9FBZ8 | Q9fbz8 streptomyc |
| 434 | 7 | 0.8 | 291 | 16 | Q8ZIM7 | Q8zim7 salmonella | 507 | 7 | 0.8 | 349 | 16 | Q8ZBG6 | Q8zbg6 yersinia pe |
| 435 | 7 | 0.8 | 291 | 16 | Q9X1N5 | Q9x1n5 thermotoga | 508 | 7 | 0.8 | 349 | 16 | Q8ZBG6 | Q8zbg6 yersinia pe |
| 436 | 7 | 0.8 | 293 | 13 | Q8AXY4 | Q8axy4 brachydanio | 509 | 7 | 0.8 | 350 | 16 | Q8FSN3 | Q8fsn3 corynebacte |
| 437 | 7 | 0.8 | 294 | 5 | Q8IIB6 | Q8iib6 plasmodium | 510 | 7 | 0.8 | 350 | 16 | Q8YNL0 | Q8ynl0 anabaena sp |
| 438 | 7 | 0.8 | 294 | 16 | Q8XF74 | Q8xf74 salmonella | 511 | 7 | 0.8 | 351 | 16 | Q9K9N0 | Q9k9n0 bacillus ha |
| 439 | 7 | 0.8 | 296 | 2 | Q8TRT1 | Q8trt1 pseudomonas | 512 | 7 | 0.8 | 351 | 16 | Q8ZDT2 | Q8zdt2 yersinia pe |
| 440 | 7 | 0.8 | 296 | 16 | Q8FRU9 | Q8fru9 corynebacte | 513 | 7 | 0.8 | 351 | 17 | Q97Z26 | Q97z26 sulfolobus |
| 441 | 7 | 0.8 | 297 | 16 | P74131 | P74131 synechocyst | 514 | 7 | 0.8 | 352 | 2 | Q8KZC4 | Q8kzc4 helicobacte |
| 442 | 7 | 0.8 | 298 | 10 | Q23367 | Q23367 arabidopsis | 515 | 7 | 0.8 | 352 | 12 | Q73518 | Q73518 potato viru |
| 443 | 7 | 0.8 | 299 | 13 | Q9DGH3 | Q9dgh3 brachydanio | 516 | 7 | 0.8 | 352 | 17 | Q9HKB6 | Q9hkb6 thermoplasm |
| 444 | 7 | 0.8 | 299 | 16 | Q55820 | Q55820 synechocyst | 517 | 7 | 0.8 | 353 | 2 | Q9S3U8 | Q9s3u8 helicobacte |
| 445 | 7 | 0.8 | 300 | 2 | Q8GC81 | Q8gc81 enterobacte | 518 | 7 | 0.8 | 353 | 2 | Q8KZ47 | Q8kz47 helicobacte |
| 446 | 7 | 0.8 | 300 | 4 | Q8W57 | Q8w57 homo sapien | 519 | 7 | 0.8 | 355 | 8 | Q47483 | Q47483 cyclus oka |
| 447 | 7 | 0.8 | 304 | 11 | Q8C7W6 | Q8c7w6 musculu | 520 | 7 | 0.8 | 355 | 12 | Q73519 | Q73519 potato viru |
| 448 | 7 | 0.8 | 304 | 16 | Q92RB7 | Q92rb7 rhizobium m | 521 | 7 | 0.8 | 355 | 12 | Q73510 | Q73510 potato viru |
| 449 | 7 | 0.8 | 305 | 4 | Q96A60 | Q96a60 homo sapien | 522 | 7 | 0.8 | 355 | 12 | Q73532 | Q73532 potato viru |
| 450 | 7 | 0.8 | 305 | 10 | Q8L4J8 | Q8l4j8 arabidopsis | 523 | 7 | 0.8 | 355 | 12 | Q73533 | Q73533 potato viru |
| 451 | 7 | 0.8 | 307 | 16 | Q8CR46 | Q8cr46 staphylococ | 524 | 7 | 0.8 | 355 | 12 | Q73515 | Q73515 potato viru |
| 452 | 7 | 0.8 | 308 | 16 | Q8NML1 | Q8nml1 corynebacte | 525 | 7 | 0.8 | 355 | 12 | Q73520 | Q73520 potato viru |
| 453 | 7 | 0.8 | 309 | 6 | Q8WNU0 | Q8wnu0 macaca fasc | 526 | 7 | 0.8 | 355 | 12 | Q73517 | Q73517 potato viru |
| 454 | 7 | 0.8 | 311 | 10 | Q9ZVU4 | Q9zvu4 arabidopsis | 527 | 7 | 0.8 | 357 | 5 | O17591 | O17591 caenorhabdi |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|----|--------|--------------------|-----|---|-----|-----|----|--------|----------------------|
| 528 | 7 | 0.8 | 358 | 16 | O05579 | O05579 mycobacteri | 601 | 7 | 0.8 | 404 | 17 | O29700 | O29700 archaeoglob |
| 529 | 7 | 0.8 | 358 | 16 | O8Y1P7 | O8Y1P7 ralstonia s | 602 | 7 | 0.8 | 406 | 5 | O8ILL1 | O8ILL1 plasmodium |
| 530 | 7 | 0.8 | 360 | 16 | O8YV68 | O8YV68 anabaena sp | 603 | 7 | 0.8 | 407 | 16 | O9CG50 | O9CG50 lactococcus |
| 531 | 7 | 0.8 | 362 | 2 | O8KZB5 | O8KZB5 helicobacte | 604 | 7 | 0.8 | 407 | 16 | O51626 | O51626 borellia bu |
| 532 | 7 | 0.8 | 362 | 2 | O9S3L9 | O9S3L9 borellia re | 605 | 7 | 0.8 | 408 | 2 | O85454 | O85454 pasteurella |
| 533 | 7 | 0.8 | 363 | 16 | O8UIK8 | O8UIK8 agrobacteri | 606 | 7 | 0.8 | 408 | 16 | O9CMP3 | O9CMP3 pasteurella |
| 534 | 7 | 0.8 | 364 | 16 | O8D549 | O8D549 vibrio vuln | 607 | 7 | 0.8 | 408 | 16 | O9RCV5 | O9RCV5 streptomyce |
| 535 | 7 | 0.8 | 365 | 10 | O8LDW7 | O8LDW7 arabidopsis | 608 | 7 | 0.8 | 408 | 16 | O8DOW6 | O8DOW6 yerinia pe |
| 536 | 7 | 0.8 | 366 | 10 | O8S052 | O8S052 oryza sativ | 609 | 7 | 0.8 | 410 | 16 | O8X9A2 | O8X9A2 escherichia |
| 537 | 7 | 0.8 | 367 | 9 | O64040 | O64040 bacterioph | 610 | 7 | 0.8 | 411 | 16 | O92X82 | O92X82 rhizobium m |
| 538 | 7 | 0.8 | 367 | 16 | O9KKV7 | O9KKV7 vibrio chol | 611 | 7 | 0.8 | 413 | 10 | O9F7S8 | O9F7S8 arabidopsis |
| 539 | 7 | 0.8 | 367 | 16 | O31982 | O31982 bacillus su | 612 | 7 | 0.8 | 414 | 13 | O980J2 | O980J2 pseudopas lae |
| 540 | 7 | 0.8 | 368 | 16 | O9A048 | O9A048 streptococ | 613 | 7 | 0.8 | 414 | 16 | O33415 | O33415 pseudomonas |
| 541 | 7 | 0.8 | 368 | 16 | O8KT74 | O8KT74 streptococ | 614 | 7 | 0.8 | 415 | 4 | O9UDU3 | O9UDU3 homo sapien |
| 542 | 7 | 0.8 | 370 | 2 | O9AE31 | O9AE31 rhizobium l | 615 | 7 | 0.8 | 415 | 16 | O8YU57 | O8YU57 anabaena sp |
| 543 | 7 | 0.8 | 370 | 17 | O974H2 | O974H2 sulfolobus | 616 | 7 | 0.8 | 418 | 11 | O63918 | O63918 mus musculu |
| 544 | 7 | 0.8 | 371 | 2 | O9RLE3 | O9RLE3 prochloroco | 617 | 7 | 0.8 | 418 | 11 | O6CBT4 | O6CBT4 mus musculu |
| 545 | 7 | 0.8 | 371 | 2 | O86250 | O86250 helicobacte | 618 | 7 | 0.8 | 422 | 16 | O8F4J2 | O8F4J2 xanthomonas |
| 546 | 7 | 0.8 | 372 | 2 | O93K39 | O93K39 klebsiella | 619 | 7 | 0.8 | 422 | 16 | O8CMQ6 | O8CMQ6 staphylococ |
| 547 | 7 | 0.8 | 375 | 2 | O9AGC8 | O9AGC8 enterobacte | 620 | 7 | 0.8 | 424 | 5 | O77237 | O77237 drosophila |
| 548 | 7 | 0.8 | 375 | 2 | O9ALY0 | O9ALY0 enterobacte | 621 | 7 | 0.8 | 426 | 16 | O9A997 | O9A997 caulobacter |
| 549 | 7 | 0.8 | 375 | 2 | O9AGC9 | O9AGC9 enterobacte | 622 | 7 | 0.8 | 429 | 5 | O811Q5 | O811Q5 plasmodium |
| 550 | 7 | 0.8 | 375 | 2 | O9AGD1 | O9AGD1 enterobacte | 623 | 7 | 0.8 | 429 | 15 | P88173 | P88173 human immu |
| 551 | 7 | 0.8 | 375 | 2 | O93T24 | O93T24 enterobacte | 624 | 7 | 0.8 | 429 | 16 | O97KC1 | O97KC1 clostridium |
| 552 | 7 | 0.8 | 375 | 16 | O9RXS7 | O9RXS7 deinococcus | 625 | 7 | 0.8 | 429 | 16 | O9Y0K6 | O9Y0K6 ralstonia s |
| 553 | 7 | 0.8 | 376 | 2 | O9AGD0 | O9AGD0 enterobacte | 626 | 7 | 0.8 | 430 | 16 | O97FJ7 | O97FJ7 clostridium |
| 554 | 7 | 0.8 | 376 | 2 | O52496 | O52496 bradyrhizob | 627 | 7 | 0.8 | 431 | 10 | O9M9S8 | O9M9S8 arabidopsis |
| 555 | 7 | 0.8 | 376 | 17 | O96VZ5 | O96VZ5 sulfolobus | 628 | 7 | 0.8 | 431 | 16 | O8G4W4 | O8G4W4 bifidobacte |
| 556 | 7 | 0.8 | 377 | 2 | O8VRL8 | O8VRL8 thermus the | 629 | 7 | 0.8 | 434 | 16 | O9L227 | O9L227 streptomyce |
| 557 | 7 | 0.8 | 377 | 16 | O92SP9 | O92SP9 rhizobium m | 630 | 7 | 0.8 | 439 | 10 | O6S915 | O6S915 arabidopsis |
| 558 | 7 | 0.8 | 379 | 16 | O9RV61 | O9RV61 deinococcus | 631 | 7 | 0.8 | 439 | 16 | O8BKU1 | O8BKU1 synchococ |
| 559 | 7 | 0.8 | 379 | 16 | O8R790 | O8R790 thermoanaer | 632 | 7 | 0.8 | 440 | 16 | O8P988 | O8P988 xanthomonas |
| 560 | 7 | 0.8 | 379 | 16 | O9RD86 | O9RD86 streptomyce | 633 | 7 | 0.8 | 441 | 2 | O9X708 | O9X708 clostridium |
| 561 | 7 | 0.8 | 380 | 8 | O48089 | O48089 lacerta viv | 634 | 7 | 0.8 | 441 | 16 | O9RSF5 | O9RSF5 deinococcus |
| 562 | 7 | 0.8 | 380 | 16 | O9LOH9 | O9LOH9 streptomyce | 635 | 7 | 0.8 | 442 | 10 | O8LME3 | O8LME3 oryza sativ |
| 563 | 7 | 0.8 | 381 | 10 | O9FJQ1 | O9FJQ1 arabidopsis | 636 | 7 | 0.8 | 442 | 16 | P96400 | P96400 mycobacteri |
| 564 | 7 | 0.8 | 381 | 16 | O8FUJ3 | O8FUJ3 corynebacte | 637 | 7 | 0.8 | 443 | 11 | O8CFY5 | O8CFY5 mus musculu |
| 565 | 7 | 0.8 | 382 | 3 | O93925 | O93925 gibberella | 638 | 7 | 0.8 | 443 | 16 | O8YVT6 | O8YVT6 anabaena sp |
| 566 | 7 | 0.8 | 382 | 16 | O25891 | O25891 helicobacte | 639 | 7 | 0.8 | 444 | 16 | O9HUR0 | O9HUR0 pseudomonas |
| 567 | 7 | 0.8 | 384 | 1 | O14094 | O14094 echinosacch | 640 | 7 | 0.8 | 446 | 16 | O8ECB3 | O8ECB3 shewanella |
| 568 | 7 | 0.8 | 384 | 10 | O9FV89 | O9FV89 petroselinu | 641 | 7 | 0.8 | 449 | 2 | O9R6Z7 | O9R6Z7 nocardioide |
| 569 | 7 | 0.8 | 384 | 10 | O9FV88 | O9FV88 petroselinu | 642 | 7 | 0.8 | 449 | 16 | O9CGJ0 | O9CGJ0 lactococcus |
| 570 | 7 | 0.8 | 385 | 12 | O8QLC1 | O8QLC1 mamestra co | 643 | 7 | 0.8 | 449 | 16 | O9WYL2 | O9WYL2 thermotoga |
| 571 | 7 | 0.8 | 385 | 16 | O92204 | O92204 rhizobium m | 644 | 7 | 0.8 | 449 | 16 | O8DFK6 | O8DFK6 vibrio vuln |
| 572 | 7 | 0.8 | 386 | 16 | O25780 | O25780 helicobacte | 645 | 7 | 0.8 | 450 | 10 | O9LPW3 | O9LPW3 arabidopsis |
| 573 | 7 | 0.8 | 386 | 16 | O9ZK50 | O9ZK50 helicobacte | 646 | 7 | 0.8 | 451 | 11 | O9D4L5 | O9D4L5 mus musculu |
| 574 | 7 | 0.8 | 387 | 5 | O8IU54 | O8IU54 plasmodium | 647 | 7 | 0.8 | 451 | 16 | O9K425 | O9K425 streptomyce |
| 575 | 7 | 0.8 | 387 | 16 | O92VH1 | O92VH1 rhizobium m | 648 | 7 | 0.8 | 452 | 2 | O9X312 | O9X312 bacillus an |
| 576 | 7 | 0.8 | 388 | 16 | O8UJL8 | O8UJL8 agrobacteri | 649 | 7 | 0.8 | 452 | 17 | O9V209 | O9V209 pyrococcus |
| 577 | 7 | 0.8 | 388 | 16 | O8UG48 | O8UG48 agrobacteri | 650 | 7 | 0.8 | 452 | 17 | O97ZD7 | O97ZD7 sulfolobus |
| 578 | 7 | 0.8 | 389 | 2 | O9AIH9 | O9AIH9 chlamydophi | 651 | 7 | 0.8 | 454 | 16 | O8FAP1 | O8FAP1 escherichia |
| 579 | 7 | 0.8 | 389 | 16 | O8P970 | O8P970 xanthomonas | 652 | 7 | 0.8 | 454 | 17 | O50131 | O50131 pyrococcus |
| 580 | 7 | 0.8 | 390 | 2 | O8KRN6 | O8KRN6 cytophaga j | 653 | 7 | 0.8 | 454 | 17 | O9VOQ7 | O9VOQ7 pyrococcus |
| 581 | 7 | 0.8 | 390 | 17 | O8TJL6 | O8TJL6 methanosarc | 654 | 7 | 0.8 | 454 | 17 | O8UL09 | O8UL09 pyrococcus |
| 582 | 7 | 0.8 | 391 | 12 | O89469 | O89469 avian pneum | 655 | 7 | 0.8 | 455 | 16 | O9ZKC8 | O9ZKC8 helicobacte |
| 583 | 7 | 0.8 | 391 | 12 | O91HN5 | O91HN5 avian pneum | 656 | 7 | 0.8 | 457 | 8 | O94QO5 | O94QO5 caiman croc |
| 584 | 7 | 0.8 | 391 | 12 | O9DKS9 | O9DKS9 avian pneum | 657 | 7 | 0.8 | 458 | 16 | O25134 | O25134 helicobacte |
| 585 | 7 | 0.8 | 391 | 16 | O8VKA2 | O8VKA2 mycobacteri | 658 | 7 | 0.8 | 460 | 4 | O9V589 | O9V589 homo sapien |
| 586 | 7 | 0.8 | 392 | 5 | O9V6N7 | O9V6N7 drosophila | 659 | 7 | 0.8 | 460 | 4 | O8TDF7 | O8TDF7 homo sapien |
| 587 | 7 | 0.8 | 393 | 10 | O8SON2 | O8SON2 oryza sativ | 660 | 7 | 0.8 | 460 | 4 | O9NWX3 | O9NWX3 homo sapien |
| 588 | 7 | 0.8 | 393 | 16 | O8ZC13 | O8ZC13 yerinia pe | 661 | 7 | 0.8 | 460 | 11 | O62739 | O62739 rattus norv |
| 589 | 7 | 0.8 | 393 | 16 | O9EWE0 | O9EWE0 streptomyce | 662 | 7 | 0.8 | 462 | 16 | O915H2 | O915H2 pseudomonas |
| 590 | 7 | 0.8 | 395 | 17 | O97VF3 | O97VF3 sulfolobus | 663 | 7 | 0.8 | 462 | 16 | O9XB12 | O9XB12 bacillus ha |
| 591 | 7 | 0.8 | 396 | 2 | O93KV6 | O93KV6 streptomyce | 664 | 7 | 0.8 | 463 | 16 | O824Y1 | O824Y1 salmonella |
| 592 | 7 | 0.8 | 396 | 2 | O93A40 | O93A40 enterococcu | 665 | 7 | 0.8 | 466 | 10 | O9LFH7 | O9LFH7 arabidopsis |
| 593 | 7 | 0.8 | 397 | 16 | O9I223 | O9I223 pseudomonas | 666 | 7 | 0.8 | 469 | 5 | O8IJV7 | O8IJV7 plasmodium |
| 594 | 7 | 0.8 | 399 | 16 | O8DGM0 | O8DGM0 synchococ | 667 | 7 | 0.8 | 471 | 16 | O9S2W9 | O9S2W9 streptomyce |
| 595 | 7 | 0.8 | 400 | 16 | O9PHA8 | O9PHA8 xylella fas | 668 | 7 | 0.8 | 472 | 2 | O8RNU1 | O8RNU1 helicobacte |
| 596 | 7 | 0.8 | 401 | 10 | O9LWM5 | O9LWM5 oryza sativ | 669 | 7 | 0.8 | 473 | 2 | O9RB20 | O9RB20 pectobacter |
| 597 | 7 | 0.8 | 401 | 16 | O9K7S7 | O9K7S7 bacillus ha | 670 | 7 | 0.8 | 473 | 16 | O8VCD4 | O8VCD4 bruceella me |
| 598 | 7 | 0.8 | 402 | 2 | O33979 | O33979 rhizobium e | 671 | 7 | 0.8 | 473 | 17 | O9YWB8 | O9YWB8 sulfolobus |
| 599 | 7 | 0.8 | 403 | 2 | O86253 | O86253 helicobacte | 672 | 7 | 0.8 | 474 | 16 | O8ZJ57 | O8ZJ57 yerinia pe |
| 600 | 7 | 0.8 | 404 | 17 | O97C00 | O97C00 thermoplas | 673 | 7 | 0.8 | 476 | 17 | O8Z2R7 | O8Z2R7 pyrobaculum |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|----|---------|---------------------|-----|---|-----|-----|----|--------|----------------------|
| 674 | 7 | 0.8 | 477 | 2 | Q59388 | Q59388 escherichia | 747 | 7 | 0.8 | 558 | 11 | Q9WTL5 | Q9wtl5 mus musculus |
| 675 | 7 | 0.8 | 477 | 16 | Q9PBD6 | Q9pbd6 xylella fas | 748 | 7 | 0.8 | 558 | 11 | Q9R1W1 | Q9r1w1 mus musculus |
| 676 | 7 | 0.8 | 478 | 2 | Q9XB65 | Q9xb65 erwinia amy | 749 | 7 | 0.8 | 558 | 16 | Q9X225 | Q9x225 thermotoga |
| 677 | 7 | 0.8 | 479 | 2 | P95598 | P95598 corynebacte | 750 | 7 | 0.8 | 559 | 16 | Q92VX5 | Q92vx5 rhizobium m |
| 678 | 7 | 0.8 | 480 | 2 | Q52080 | Q52080 pseudomonas | 751 | 7 | 0.8 | 561 | 5 | Q9VZM5 | Q9vzm5 drosophila |
| 679 | 7 | 0.8 | 480 | 4 | Q8ND99 | Q8nd99 homo sapien | 752 | 7 | 0.8 | 564 | 3 | Q8JIV3 | Q8jiv3 drosophila |
| 680 | 7 | 0.8 | 480 | 16 | Q8DLQ3 | Q8dlq3 synchococc | 753 | 7 | 0.8 | 568 | 16 | Q8DM73 | Q8dm73 synchococc |
| 681 | 7 | 0.8 | 482 | 5 | Q9UA14 | Q9ua14 dictyosteli | 754 | 7 | 0.8 | 575 | 10 | Q8LGJ2 | Q8lgj2 arabidopsis |
| 682 | 7 | 0.8 | 482 | 16 | Q8UA11 | Q8ua11 agrobacteri | 755 | 7 | 0.8 | 575 | 10 | Q9LZM8 | Q9lzm8 arabidopsis |
| 683 | 7 | 0.8 | 482 | 17 | Q54610 | Q54610 halobacteri | 756 | 7 | 0.8 | 575 | 10 | Q8H147 | Q8h147 arabidopsis |
| 684 | 7 | 0.8 | 483 | 5 | Q9Y0A6 | Q9y0a6 dictyosteli | 757 | 7 | 0.8 | 575 | 16 | Q8FW30 | Q8fw30 brucella su |
| 685 | 7 | 0.8 | 483 | 5 | Q8IJCO | Q8ijco plasmodium | 758 | 7 | 0.8 | 576 | 5 | Q8ID23 | Q8id23 plasmodium |
| 686 | 7 | 0.8 | 483 | 5 | Q8IASN5 | Q8iasn5 caulobacter | 759 | 7 | 0.8 | 576 | 16 | P71664 | P71664 mycobacteri |
| 687 | 7 | 0.8 | 484 | 16 | Q9KRH2 | Q9krh2 vibrio chol | 760 | 7 | 0.8 | 577 | 2 | Q9L2F3 | Q9l2f3 streptomyce |
| 688 | 7 | 0.8 | 485 | 5 | O18125 | O18125 caenorhabdi | 761 | 7 | 0.8 | 580 | 2 | Q8KW93 | Q8kw93 ruegeria sp |
| 689 | 7 | 0.8 | 488 | 17 | Q57576 | Q57576 methanococc | 762 | 7 | 0.8 | 581 | 5 | Q23969 | Q23969 drosophila |
| 690 | 7 | 0.8 | 489 | 16 | Q8XQL1 | Q8xql1 ralstonia s | 763 | 7 | 0.8 | 581 | 10 | Q8RX72 | Q8rx72 arabidopsis |
| 691 | 7 | 0.8 | 491 | 10 | Q9FHE2 | Q9fhe2 arabidopsis | 764 | 7 | 0.8 | 583 | 11 | Q8CEW7 | Q8cew7 mus musculus |
| 692 | 7 | 0.8 | 491 | 17 | Q8TWJ2 | Q8tmj2 methanocarc | 765 | 7 | 0.8 | 584 | 16 | Q8KCC2 | Q8kcc2 chlorobium |
| 693 | 7 | 0.8 | 493 | 16 | Q8XD80 | Q8xd80 escherichia | 766 | 7 | 0.8 | 587 | 5 | Q8IT76 | Q8it76 manduca sex |
| 694 | 7 | 0.8 | 497 | 16 | Q98I38 | Q98i38 rhizobium l | 767 | 7 | 0.8 | 587 | 16 | Q9KR30 | Q9kr30 vibrio chol |
| 695 | 7 | 0.8 | 498 | 5 | Q9XU88 | Q9xu88 caenorhabdi | 768 | 7 | 0.8 | 588 | 2 | Q8G9X8 | Q8g9x8 escherichia |
| 696 | 7 | 0.8 | 498 | 16 | Q8FVY2 | Q8fvy2 brucella su | 769 | 7 | 0.8 | 588 | 10 | Q9MA86 | Q9ma86 arabidopsis |
| 697 | 7 | 0.8 | 500 | 16 | Q9KB49 | Q9kb49 bacillus ha | 770 | 7 | 0.8 | 589 | 5 | Q94216 | Q94216 caenorhabdi |
| 698 | 7 | 0.8 | 500 | 16 | Q8REU1 | Q8reu1 fusobacteri | 771 | 7 | 0.8 | 589 | 11 | Q8VDL5 | Q8vdl5 mus musculus |
| 699 | 7 | 0.8 | 500 | 16 | Q8F9S7 | Q8f9s7 leptospira | 772 | 7 | 0.8 | 589 | 16 | Q92HC1 | Q92hc1 rickettsia |
| 700 | 7 | 0.8 | 504 | 17 | Q978P9 | Q978p9 thermoplasma | 773 | 7 | 0.8 | 590 | 10 | Q8S4W7 | Q8s4w7 vitis vinif |
| 701 | 7 | 0.8 | 504 | 16 | Q8DB29 | Q8db29 vibrio vuln | 774 | 7 | 0.8 | 592 | 4 | Q8NC78 | Q8nc78 homo sapien |
| 702 | 7 | 0.8 | 505 | 10 | O65558 | O65558 arabidopsis | 775 | 7 | 0.8 | 595 | 5 | O45934 | O45934 caenorhabdi |
| 703 | 7 | 0.8 | 505 | 10 | O48645 | O48645 lycopersico | 776 | 7 | 0.8 | 600 | 16 | Q9KBX2 | Q9kbx2 bacillus ha |
| 704 | 7 | 0.8 | 507 | 4 | Q9P080 | Q9p080 homo sapien | 777 | 7 | 0.8 | 605 | 3 | Q96VA3 | Q96va3 aspergillus |
| 705 | 7 | 0.8 | 507 | 11 | Q9Z125 | Q9z125 mus musculus | 778 | 7 | 0.8 | 605 | 16 | Q8U7N6 | Q8u7n6 agrobacteri |
| 706 | 7 | 0.8 | 507 | 12 | Q9E1X3 | Q9e1x3 cercothec | 779 | 7 | 0.8 | 608 | 16 | Q9PR80 | Q9pr80 ureaplasma |
| 707 | 7 | 0.8 | 509 | 2 | Q9AN21 | Q9an21 bradyrhizob | 780 | 7 | 0.8 | 609 | 16 | Q9P193 | Q9p193 campylobact |
| 708 | 7 | 0.8 | 510 | 10 | Q9LXK1 | Q9lex1 arabidopsis | 781 | 7 | 0.8 | 610 | 2 | Q9S0M8 | Q9s0m8 deinococcus |
| 709 | 7 | 0.8 | 512 | 10 | Q8LSK3 | Q8lsk3 physcomitre | 782 | 7 | 0.8 | 613 | 5 | Q8SWS5 | Q8sws5 drosophila |
| 710 | 7 | 0.8 | 512 | 16 | Q8NM17 | Q8nm17 rhizobium l | 783 | 7 | 0.8 | 616 | 16 | Q8UAV6 | Q8uav6 agrobacteri |
| 711 | 7 | 0.8 | 512 | 16 | P71828 | P71828 mycobacteri | 784 | 7 | 0.8 | 617 | 2 | Q9EVD8 | Q9evd8 pseudomonas |
| 712 | 7 | 0.8 | 513 | 16 | Q8FMH3 | Q8fmh3 corynebacte | 785 | 7 | 0.8 | 617 | 2 | Q9EVB2 | Q9evb2 pseudomonas |
| 713 | 7 | 0.8 | 517 | 16 | Q9RDI1 | Q9rdi1 streptomyce | 786 | 7 | 0.8 | 617 | 16 | Q9RUN0 | Q9run0 deinococcus |
| 714 | 7 | 0.8 | 517 | 17 | Q97BE6 | Q97be6 thermoplasma | 787 | 7 | 0.8 | 620 | 11 | Q9WYB6 | Q9wyb6 thermotoga |
| 715 | 7 | 0.8 | 519 | 4 | Q96BA8 | Q96ba8 homo sapien | 788 | 7 | 0.8 | 621 | 11 | Q8CAA7 | Q8caa7 mus musculus |
| 716 | 7 | 0.8 | 520 | 11 | Q91W70 | Q91w70 mus musculus | 789 | 7 | 0.8 | 621 | 16 | Q8EUH7 | Q8euh7 mycoplasma |
| 717 | 7 | 0.8 | 521 | 11 | Q8CBN2 | Q8cbn2 mus musculus | 790 | 7 | 0.8 | 624 | 13 | Q90WU2 | Q90wu2 fundulus he |
| 718 | 7 | 0.8 | 521 | 11 | Q8CH61 | Q8ch61 mus musculus | 791 | 7 | 0.8 | 630 | 5 | Q9VP21 | Q9vp21 drosophila |
| 719 | 7 | 0.8 | 521 | 11 | Q8BH52 | Q8bh52 mus musculus | 792 | 7 | 0.8 | 632 | 16 | Q928V3 | Q928v3 listeria in |
| 720 | 7 | 0.8 | 525 | 16 | Q9A7F8 | Q9a7f8 caulobacter | 793 | 7 | 0.8 | 632 | 16 | Q8Y4U6 | Q8y4u6 listeria mo |
| 721 | 7 | 0.8 | 527 | 13 | Q8AXY3 | Q8axy3 fugu rubrip | 794 | 7 | 0.8 | 634 | 5 | Q9VRN6 | Q9vrn6 drosophila |
| 722 | 7 | 0.8 | 527 | 13 | Q8AXY2 | Q8axy2 tetraodon n | 795 | 7 | 0.8 | 634 | 5 | Q95PL8 | Q95pl8 trypanosoma |
| 723 | 7 | 0.8 | 529 | 16 | Q8ET94 | Q8et94 oceanobacil | 796 | 7 | 0.8 | 634 | 17 | Q8PRY4 | Q8pry4 methanocarc |
| 724 | 7 | 0.8 | 530 | 5 | O01983 | O01983 caenorhabdi | 797 | 7 | 0.8 | 635 | 16 | Q9GQP3 | Q9gqp3 mycoplasma |
| 725 | 7 | 0.8 | 531 | 12 | Q9E3S5 | Q9e3s5 human parvo | 798 | 7 | 0.8 | 636 | 4 | Q8N7M8 | Q8n7m8 homo sapien |
| 726 | 7 | 0.8 | 532 | 11 | Q9R1W2 | Q9r1w2 mus musculus | 799 | 7 | 0.8 | 641 | 4 | Q9BXI0 | Q9bx10 homo sapien |
| 727 | 7 | 0.8 | 532 | 11 | Q9WTL6 | Q9wtl6 mus musculus | 800 | 7 | 0.8 | 641 | 7 | Q8SNA1 | Q8sna1 homo sapien |
| 728 | 7 | 0.8 | 534 | 17 | Q9YDN5 | Q9ydn5 aeropyrum p | 801 | 7 | 0.8 | 641 | 7 | Q8MHQ1 | Q8mhq1 homo sapien |
| 729 | 7 | 0.8 | 536 | 16 | Q8FND6 | Q8fnd6 corynebacte | 802 | 7 | 0.8 | 642 | 17 | Q8TIE9 | Q8tie9 methanocarc |
| 730 | 7 | 0.8 | 537 | 5 | Q8IPG6 | Q8ipg6 drosophila | 803 | 7 | 0.8 | 642 | 17 | Q8PXV1 | Q8pxv1 methanocarc |
| 731 | 7 | 0.8 | 537 | 16 | Q8X542 | Q8x542 ralstonia s | 804 | 7 | 0.8 | 643 | 16 | Q8G292 | Q8g292 brucella su |
| 732 | 7 | 0.8 | 537 | 16 | Q9AK40 | Q9ak40 streptomyce | 805 | 7 | 0.8 | 646 | 5 | Q9XU22 | Q9xu22 caenorhabdi |
| 733 | 7 | 0.8 | 539 | 16 | Q9ADA6 | Q9ada6 streptomyce | 806 | 7 | 0.8 | 652 | 16 | Q98IC4 | Q98ic4 rhizobium l |
| 734 | 7 | 0.8 | 539 | 3 | Q96U52 | Q96u52 neurospora | 807 | 7 | 0.8 | 654 | 16 | Q9CHI7 | Q9chi7 lactococcus |
| 735 | 7 | 0.8 | 543 | 5 | O18556 | O18556 musca domes | 808 | 7 | 0.8 | 655 | 16 | Q69855 | Q69855 streptomyce |
| 736 | 7 | 0.8 | 544 | 11 | Q8BW24 | Q8bw24 mus musculus | 809 | 7 | 0.8 | 658 | 2 | Q8VP71 | Q8vp71 klebsiella |
| 737 | 7 | 0.8 | 546 | 11 | Q8BIL1 | Q8bil1 mus musculus | 810 | 7 | 0.8 | 658 | 10 | Q9CAN3 | Q9can3 arabidopsis |
| 738 | 7 | 0.8 | 548 | 16 | Q92NZ9 | Q92nz9 rhizobium m | 811 | 7 | 0.8 | 659 | 2 | Q8GH79 | Q8gh79 chlamydomphi |
| 739 | 7 | 0.8 | 548 | 16 | Q9EWY9 | Q9ewy9 streptomyce | 812 | 7 | 0.8 | 659 | 16 | Q8PFG6 | Q8pfg6 xanthomonas |
| 740 | 7 | 0.8 | 550 | 16 | Q98IS1 | Q98is1 rhizobium l | 813 | 7 | 0.8 | 659 | 16 | Q8P5D9 | Q8p5d9 xanthomonas |
| 741 | 7 | 0.8 | 551 | 5 | Q9GYC4 | Q9gyc4 leishmania | 814 | 7 | 0.8 | 662 | 5 | O44832 | O44832 caenorhabdi |
| 742 | 7 | 0.8 | 551 | 16 | Q8NS49 | Q8ns49 corynebacte | 815 | 7 | 0.8 | 670 | 16 | Q8EOX1 | Q8eox1 streptococc |
| 743 | 7 | 0.8 | 551 | 16 | Q8FR55 | Q8fr55 corynebacte | 816 | 7 | 0.8 | 671 | 16 | Q9KVH9 | Q9kvh9 vibrio chol |
| 744 | 7 | 0.8 | 553 | 2 | Q8KLJ4 | Q8klj4 streptomyce | 817 | 7 | 0.8 | 674 | 4 | Q9HAB3 | Q9hab3 homo sapien |
| 745 | 7 | 0.8 | 554 | 13 | Q8AUN9 | Q8aun9 fugu rubrip | 818 | 7 | 0.8 | 674 | 11 | Q8BZF7 | Q8bzif7 mus musculus |
| 746 | 7 | 0.8 | 557 | 16 | Q8GLJ5 | Q8glj5 brucella su | 819 | 7 | 0.8 | 676 | 16 | Q8YVF5 | Q8yvf5 anabaena sp |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|----|---------|---------------------|-----|---|-----|------|----|--------|---------------------|
| 820 | 7 | 0.8 | 678 | 2 | Q54679 | Q54679 streptomyce | 893 | 7 | 0.8 | 829 | 2 | Q9F9G0 | Q9F9G0 helicobacte |
| 821 | 7 | 0.8 | 679 | 5 | Q9U0V8 | Q9U0V8 leishmania | 894 | 7 | 0.8 | 833 | 5 | Q9VLM1 | Q9VLM1 drosophila |
| 822 | 7 | 0.8 | 680 | 16 | Q9XD18 | Q9XD18 bacteroides | 895 | 7 | 0.8 | 836 | 17 | Q8U4M8 | Q8U4M8 pyrococcus |
| 823 | 7 | 0.8 | 681 | 16 | Q927H1 | Q927H1 listeria in | 896 | 7 | 0.8 | 839 | 5 | Q93Q01 | Q93Q01 helicobacte |
| 824 | 7 | 0.8 | 682 | 16 | Q8Y410 | Q8Y410 listeria mo | 897 | 7 | 0.8 | 839 | 5 | Q9NHN7 | Q9NHN7 drosophila |
| 825 | 7 | 0.8 | 683 | 4 | Q8TC07 | Q8TC07 homo sapien | 898 | 7 | 0.8 | 839 | 5 | Q9Y103 | Q9Y103 drosophila |
| 826 | 7 | 0.8 | 684 | 17 | Q8PVG2 | Q8PVG2 methanosarc | 899 | 7 | 0.8 | 843 | 5 | Q9VY15 | Q9VY15 drosophila |
| 827 | 7 | 0.8 | 685 | 5 | Q21649 | Q21649 caenorhabdi | 900 | 7 | 0.8 | 850 | 16 | Q66691 | Q66691 aquifex aeo |
| 828 | 7 | 0.8 | 686 | 16 | Q8XW44 | Q8XW44 ralestonia s | 901 | 7 | 0.8 | 855 | 17 | Q57819 | Q57819 pyrococcus |
| 829 | 7 | 0.8 | 701 | 4 | Q96K51 | Q96K51 homo sapien | 902 | 7 | 0.8 | 855 | 17 | Q9V2M1 | Q9V2M1 pyrococcus |
| 830 | 7 | 0.8 | 708 | 10 | Q8S555 | Q8S555 picea abies | 903 | 7 | 0.8 | 861 | 2 | Q9F9G2 | Q9F9G2 helicobacte |
| 831 | 7 | 0.8 | 708 | 12 | Q9DKW1 | Q9DKW1 walrus cali | 904 | 7 | 0.8 | 861 | 2 | Q9F9F7 | Q9F9F7 helicobacte |
| 832 | 7 | 0.8 | 711 | 16 | Q9G9S17 | Q9G9S17 staphylococ | 905 | 7 | 0.8 | 861 | 2 | Q9F9G5 | Q9F9G5 helicobacte |
| 833 | 7 | 0.8 | 718 | 5 | Q9B107 | Q9B107 entamoeba h | 906 | 7 | 0.8 | 862 | 2 | Q9F9G3 | Q9F9G3 helicobacte |
| 834 | 7 | 0.8 | 722 | 16 | Q9F322 | Q9F322 streptomyce | 907 | 7 | 0.8 | 862 | 10 | Q9FTQ0 | Q9FTQ0 oryza sativ |
| 835 | 7 | 0.8 | 726 | 16 | Q8XR29 | Q8XR29 ralestonia s | 908 | 7 | 0.8 | 864 | 2 | Q9F9G1 | Q9F9G1 helicobacte |
| 836 | 7 | 0.8 | 726 | 17 | Q9YE25 | Q9YE25 aeropyrum p | 909 | 7 | 0.8 | 864 | 2 | Q93Q00 | Q93Q00 helicobacte |
| 837 | 7 | 0.8 | 735 | 16 | P74453 | P74453 synechocyst | 910 | 7 | 0.8 | 866 | 2 | Q9F9G6 | Q9F9G6 helicobacte |
| 838 | 7 | 0.8 | 739 | 10 | Q9LWY3 | Q9LWY3 oryza sativ | 911 | 7 | 0.8 | 866 | 5 | Q9VF20 | Q9VF20 drosophila |
| 839 | 7 | 0.8 | 739 | 16 | Q8FMZ8 | Q8FMZ8 corynebacte | 912 | 7 | 0.8 | 872 | 4 | Q9H3R1 | Q9H3R1 homo sapien |
| 840 | 7 | 0.8 | 745 | 16 | Q8DHM6 | Q8DHM6 synechococc | 913 | 7 | 0.8 | 872 | 11 | Q9EQW8 | Q9EQW8 mus musculu |
| 841 | 7 | 0.8 | 746 | 16 | Q6G902 | Q6G902 aquifex aeo | 914 | 7 | 0.8 | 877 | 16 | Q9RK68 | Q9RK68 streptomyce |
| 842 | 7 | 0.8 | 755 | 5 | Q9VC61 | Q9VC61 drosophila | 915 | 7 | 0.8 | 879 | 2 | Q8KKY1 | Q8KKY1 rhizobium e |
| 843 | 7 | 0.8 | 756 | 16 | Q9A888 | Q9A888 caulobacter | 916 | 7 | 0.8 | 880 | 16 | Q8FSG7 | Q8FSG7 corynebacte |
| 844 | 7 | 0.8 | 757 | 16 | Q914S8 | Q914S8 pseudomonas | 917 | 7 | 0.8 | 888 | 5 | Q9VDD8 | Q9VDD8 drosophila |
| 845 | 7 | 0.8 | 759 | 5 | Q8IT91 | Q8IT91 ancyllostoma | 918 | 7 | 0.8 | 893 | 4 | Q9NZ12 | Q9NZ12 homo sapien |
| 846 | 7 | 0.8 | 760 | 12 | Q9PZT8 | Q9PZT8 human parvo | 919 | 7 | 0.8 | 893 | 4 | Q8TEA7 | Q8TEA7 homo sapien |
| 847 | 7 | 0.8 | 761 | 12 | Q9PZU0 | Q9PZU0 human parvo | 920 | 7 | 0.8 | 896 | 5 | Q8MQX4 | Q8MQX4 drosophila |
| 848 | 7 | 0.8 | 763 | 10 | Q9SNP3 | Q9SNP3 oryza sativ | 921 | 7 | 0.8 | 899 | 5 | Q9NFX2 | Q9NFX2 drosophila |
| 849 | 7 | 0.8 | 764 | 10 | Q49371 | Q49371 arabidopsis | 922 | 7 | 0.8 | 899 | 5 | Q9VGZ3 | Q9VGZ3 drosophila |
| 850 | 7 | 0.8 | 765 | 3 | Q9UVY1 | Q9UVY1 pneumocysti | 923 | 7 | 0.8 | 899 | 5 | Q76935 | Q76935 drosophila |
| 851 | 7 | 0.8 | 765 | 10 | Q949H5 | Q949H5 fragaria an | 924 | 7 | 0.8 | 899 | 10 | Q9C895 | Q9C895 arabidopsis |
| 852 | 7 | 0.8 | 765 | 12 | Q9PZT6 | Q9PZT6 human parvo | 925 | 7 | 0.8 | 900 | 3 | Q74925 | Q74925 schizosacch |
| 853 | 7 | 0.8 | 767 | 16 | Q921B3 | Q921B3 rickettsia | 926 | 7 | 0.8 | 900 | 16 | Q8YZU3 | Q8YZU3 anabaena sp |
| 854 | 7 | 0.8 | 773 | 12 | Q913X1 | Q913X1 human parvo | 927 | 7 | 0.8 | 908 | 11 | Q8C0J0 | Q8C0J0 mus musculu |
| 855 | 7 | 0.8 | 773 | 12 | Q913W8 | Q913W8 human parvo | 928 | 7 | 0.8 | 908 | 11 | Q8BMI4 | Q8BMI4 mus musculu |
| 856 | 7 | 0.8 | 781 | 12 | Q8JN56 | Q8JN56 human parvo | 929 | 7 | 0.8 | 912 | 16 | Q8P957 | Q8P957 xanthomonas |
| 857 | 7 | 0.8 | 781 | 12 | Q65789 | Q65789 human parvo | 930 | 7 | 0.8 | 928 | 11 | Q8BW71 | Q8BW71 mus musculu |
| 858 | 7 | 0.8 | 781 | 12 | P90223 | P90223 human parvo | 931 | 7 | 0.8 | 929 | 5 | Q9VR32 | Q9VR32 drosophila |
| 859 | 7 | 0.8 | 781 | 12 | Q85191 | Q85191 human parvo | 932 | 7 | 0.8 | 929 | 5 | Q95X17 | Q95X17 caenorhabdi |
| 860 | 7 | 0.8 | 781 | 12 | P90221 | P90221 human parvo | 933 | 7 | 0.8 | 936 | 5 | Q95VU2 | Q95VU2 homarus ame |
| 861 | 7 | 0.8 | 781 | 12 | P89318 | P89318 human parvo | 934 | 7 | 0.8 | 941 | 16 | Q9Z1W5 | Q9Z1W5 streptomyce |
| 862 | 7 | 0.8 | 781 | 12 | Q8JVE3 | Q8JVE3 erythroviru | 935 | 7 | 0.8 | 942 | 11 | Q8BG22 | Q8BG22 mus musculu |
| 863 | 7 | 0.8 | 781 | 12 | P90224 | P90224 human parvo | 936 | 7 | 0.8 | 946 | 5 | Q9U6B5 | Q9U6B5 dictyosteli |
| 864 | 7 | 0.8 | 781 | 12 | Q9PZT0 | Q9PZT0 human parvo | 937 | 7 | 0.8 | 960 | 16 | Q8Y377 | Q8Y377 ralestonia s |
| 865 | 7 | 0.8 | 781 | 12 | P89316 | P89316 human parvo | 938 | 7 | 0.8 | 962 | 10 | P93425 | P93425 oryza sativ |
| 866 | 7 | 0.8 | 781 | 12 | P89320 | P89320 human parvo | 939 | 7 | 0.8 | 977 | 5 | Q9W3Y4 | Q9W3Y4 drosophila |
| 867 | 7 | 0.8 | 781 | 12 | Q85117 | Q85117 human parvo | 940 | 7 | 0.8 | 982 | 5 | Q9VUA5 | Q9VUA5 drosophila |
| 868 | 7 | 0.8 | 781 | 12 | P89321 | P89321 human parvo | 941 | 7 | 0.8 | 991 | 11 | Q8BX56 | Q8BX56 mus musculu |
| 869 | 7 | 0.8 | 781 | 12 | Q9UGS0 | Q9UGS0 human parvo | 942 | 7 | 0.8 | 993 | 2 | Q8KUA3 | Q8KUA3 enterococcu |
| 870 | 7 | 0.8 | 781 | 12 | P90222 | P90222 human parvo | 943 | 7 | 0.8 | 993 | 2 | Q52054 | Q52054 enterococcu |
| 871 | 7 | 0.8 | 781 | 12 | P90228 | P90228 human eryth | 944 | 7 | 0.8 | 1000 | 5 | Q96696 | Q96696 heliothis v |
| 872 | 7 | 0.8 | 781 | 12 | Q9WKQ0 | Q9WKQ0 human parvo | 945 | 7 | 0.8 | 1002 | 15 | Q8JC77 | Q8JC77 human immun |
| 873 | 7 | 0.8 | 781 | 12 | P89319 | P89319 human parvo | 946 | 7 | 0.8 | 1005 | 15 | Q8JBR1 | Q8JBR1 human immun |
| 874 | 7 | 0.8 | 781 | 12 | Q90200 | Q90200 human parvo | 947 | 7 | 0.8 | 1008 | 5 | Q8MR76 | Q8MR76 drosophila |
| 875 | 7 | 0.8 | 781 | 12 | Q8JYD9 | Q8JYD9 erythroviru | 948 | 7 | 0.8 | 1014 | 16 | Q9JZG4 | Q9JZG4 neisseria m |
| 876 | 7 | 0.8 | 781 | 12 | Q9JGP8 | Q9JGP8 human parvo | 949 | 7 | 0.8 | 1014 | 16 | Q9JUK9 | Q9JUK9 neisseria m |
| 877 | 7 | 0.8 | 781 | 12 | P89317 | P89317 human parvo | 950 | 7 | 0.8 | 1016 | 4 | Q9Z585 | Q9Z585 homo sapien |
| 878 | 7 | 0.8 | 783 | 10 | Q9C5L5 | Q9C5L5 arabidopsis | 951 | 7 | 0.8 | 1018 | 16 | Q981G7 | Q981G7 rhizobium l |
| 879 | 7 | 0.8 | 783 | 10 | Q8H1E5 | Q8H1E5 arabidopsis | 952 | 7 | 0.8 | 1020 | 5 | Q95XU2 | Q95XU2 caenorhabdi |
| 880 | 7 | 0.8 | 783 | 16 | Q99Q04 | Q99Q04 streptomyce | 953 | 7 | 0.8 | 1022 | 15 | Q90PW5 | Q90PW5 chimpanzee |
| 881 | 7 | 0.8 | 790 | 16 | Q8YSA9 | Q8YSA9 anabaena sp | 954 | 7 | 0.8 | 1024 | 5 | Q8MRZ8 | Q8MRZ8 drosophila |
| 882 | 7 | 0.8 | 806 | 16 | Q92YQ6 | Q92YQ6 rhizobium m | 955 | 7 | 0.8 | 1026 | 16 | Q8EZW3 | Q8EZW3 leptospira |
| 883 | 7 | 0.8 | 808 | 5 | Q8T5X1 | Q8T5X1 anopheles g | 956 | 7 | 0.8 | 1030 | 3 | Q12180 | Q12180 saccharomyc |
| 884 | 7 | 0.8 | 808 | 17 | Q58796 | Q58796 methanococc | 957 | 7 | 0.8 | 1034 | 16 | Q8FD25 | Q8FD25 escherichia |
| 885 | 7 | 0.8 | 815 | 16 | P96644 | P96644 bacillus su | 958 | 7 | 0.8 | 1036 | 4 | Q9NZV1 | Q9NZV1 homo sapien |
| 886 | 7 | 0.8 | 815 | 15 | Q994H2 | Q994H2 human immun | 959 | 7 | 0.8 | 1038 | 13 | Q90784 | Q90784 gallus gall |
| 887 | 7 | 0.8 | 818 | 5 | Q9VS55 | Q9VS55 drosophila | 960 | 7 | 0.8 | 1043 | 11 | Q91Z79 | Q91Z79 rattus norv |
| 888 | 7 | 0.8 | 818 | 5 | Q8TA44 | Q8TA44 drosophila | 961 | 7 | 0.8 | 1056 | 5 | Q9W3H0 | Q9W3H0 drosophila |
| 889 | 7 | 0.8 | 822 | 4 | Q96GV6 | Q96GV6 homo sapien | 962 | 7 | 0.8 | 1059 | 5 | Q9U720 | Q9U720 dictyosteli |
| 890 | 7 | 0.8 | 824 | 11 | Q8BXH3 | Q8BXH3 mus musculu | 963 | 7 | 0.8 | 1072 | 12 | Q85672 | Q85672 diadromus p |
| 891 | 7 | 0.8 | 826 | 2 | Q9F9G4 | Q9F9G4 helicobacte | 964 | 7 | 0.8 | 1072 | 16 | Q92X83 | Q92X83 rhizobium m |
| 892 | 7 | 0.8 | 829 | 2 | Q9F9F8 | Q9F9F8 helicobacte | 965 | 7 | 0.8 | 1077 | 13 | Q8UWA6 | Q8UWA6 tribolodon |

966 7 0.8 1080 13 Q8AVN3
 967 7 0.8 1082 2 Q9R964
 968 7 0.8 1083 4 Q15386
 969 7 0.8 1083 11 Q8CDJ1
 970 7 0.8 1083 11 Q8BQ26
 971 7 0.8 1085 5 Q81521
 972 7 0.8 1088 13 Q8AVH7
 973 7 0.8 1092 16 Q8EUS1
 974 7 0.8 1100 12 Q91LA3
 975 7 0.8 1109 2 Q53505
 976 7 0.8 1109 2 Q8KY00
 977 7 0.8 1112 2 Q52781
 978 7 0.8 1117 10 Q9LT22
 979 7 0.8 1122 16 Q8FW81
 980 7 0.8 1131 16 Q9RV56
 981 7 0.8 1138 5 Q8IS12
 982 7 0.8 1140 17 Q9HK21
 983 7 0.8 1146 3 Q6685
 984 7 0.8 1149 17 Q8TP29
 985 7 0.8 1155 10 Q8S869
 986 7 0.8 1155 12 Q9QW4
 987 7 0.8 1155 12 Q9Q816
 988 7 0.8 1160 2 Q8RNV8
 989 7 0.8 1163 2 Q45850
 990 7 0.8 1165 16 Q98LX9
 991 7 0.8 1170 16 Q8YC44
 992 7 0.8 1174 5 Q8WR4
 993 7 0.8 1180 17 Q8TI57
 994 7 0.8 1183 16 Q8PL49
 995 7 0.8 1185 4 Q13135
 996 7 0.8 1191 10 Q9C873
 997 7 0.8 1202 4 Q8N412
 998 7 0.8 1202 4 Q13136
 999 7 0.8 1204 5 Q93161
 1000 7 0.8 1208 12 Q39273

ALIGNMENTS

RESULT 1
 Q93GI2 PRELIMINARY; PRT; 927 AA.
 ID Q93GI2 PRELIMINARY; PRT; 927 AA.
 AC Q93GI2
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE RTX toxin.
 GN MBXA.
 OS Moraxella bovis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Moraxellaceae; Moraxella.
 OX NCBI TaxID=476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Tifton 1;
 RX MEDLINE=21388402; PubMed=11497442;
 RA Angelos J.A., Hess J.F., George L.W.;
 RT "Cloning and characterization of a Moraxella bovis cytotoxin gene."
 RL Am. J. Vet. Res. 62:1222-1228 (2001).
 DR EMBL; AF205359; AAK84651.1;
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR011343; Hemlysen_Ca_bind.
 DR InterPro; IPR003995; RTX.
 DR Pfam; PF00353; hemolysinCbind; 5.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PR00313; CAENDNGRPT.
 DR PRINTS; PR01488; RTXTOXINA.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00330; HEMOLYSIN CALCIIUM; 1.
 SQ SEQUENCE 927 AA; 98845 MW; F4B703577E10A96D CRC64;
 Query Match 89.1%; Score 826; DB 2; Length 927;

Best Local Similarity 99.9%; Pred. No. 0;
 Matches 926; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MSNINVKSIQAGLNSTKGLKXLYLAIPKDYDPQKGTGTLNDFIKAADELGIARLAEEP 60
 Db 1 MSNINVKSIQAGLNSTKGLKXLYLAIPKDYDPQKGTGTLNDFIKAADELGIARLAEEP 60
 Qy 61 NHTETAKSVDTVNOFLSLTQTGIAISATKLEFLOKHSNKLAKGLDSVENIDRKLGA 120
 Db 61 NHTETAKSVDTVNOFLSLTQTGIAISATKLEFLOKHSNKLAKGLDSVENIDRKLGA 120
 Qy 121 SNVLSTLSSFLGTALAGIELDSLIIKGDAAADALAKASIDLINEIIGNLSQSQTIEAFS 180
 Db 121 SNVLSTLSSFLGTALAGIELDSLIIKGDAAADALAKASIDLINEIIGNLSQSQTIEAFS 180
 Qy 181 SQAAGSTTSQAAGFSNIGNKLNFSKTNLGLIITGLSGISAGFALADKNASTGK 240
 Db 181 SQAAGSTTSQAAGFSNIGNKLNFSKTNLGLIITGLSGISAGFALADKNASTGK 240
 Qy 241 KVAAGFELSNOVIGNVTKATSSVYLAQORVAAGLSTTGAAVALITSSITMLAISPLAFNAA 300
 Db 241 KVAAGFELSNOVIGNVTKATSSVYLAQORVAAGLSTTGAAVALITSSITMLAISPLAFNAA 300
 Qy 301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQGVGTIEASLTITSTALGAVSAGVSAAV 360
 Db 301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQGVGTIEASLTITSTALGAVSAGVSAAV 360
 Qy 361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLOKILEWEKQNGQNYFDKGYD 420
 Db 361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLOKILEWEKQNGQNYFDKGYD 420
 Qy 421 SRVAAVLANNLKFLSELNKELEAEVTAITQQRWDNNIGELAGITKLGRIKSGKAYADA 480
 Db 421 SRVAAVLANNLKFLSELNKELEAEVTAITQQRWDNNIGELAGITKLGRIKSGKAYADA 480
 Qy 481 FEDGKVEAGSNITLDAKTGIIIDISNSNGKKTQALHFTSPLLTAGTESRRLTNGKYSYI 540
 Db 481 FEDGKVEAGSNITLDAKTGIIIDISNSNGKKTQALHFTSPLLTAGTESRRLTNGKYSYI 540
 Qy 541 NKLKFGRVKQWQVTDGEASSKLDPSKVIQVATEGTDEIGLIIVNAKAGNDDIPVGGQKM 600
 Db 541 NKLKFGRVKQWQVTDGEASSKLDPSKVIQVATEGTDEIGLIIVNAKAGNDDIPVGGQKM 600
 Qy 601 NIDGGGHDHVRVFKSGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKQETKVGK 660
 Db 601 NIDGGGHDHVRVFKSGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKQETKVGK 660
 Qy 661 RTETIOYRDYELRKRVGYQSTDNLKSVEEVIQSFNDVFKGSKFNDI FHSGEGLDLDG 720
 Db 661 RTETIOYRDYELRKRVGYQSTDNLKSVEEVIQSFNDVFKGSKFNDI FHSGEGLDLDG 720
 Qy 721 GAGDRLFGKGNDRLSGDEGDLDDGGSGDDVNLGGAGNDVYIFRKGNDNTLYDGTGN 780
 Db 721 GAGDRLFGKGNDRLSGDEGDLDDGGSGDDVNLGGAGNDVYIFRKGNDNTLYDGTGN 780
 Qy 781 DKLAFADANISDMIERTEKGIIVKRNHDSGSIIPRWIITSNLQYQSNKTDHKBOLI 840
 Db 781 DKLAFADANISDMIERTEKGIIVKRNHDSGSIIPRWIITSNLQYQSNKTDHKBOLI 840
 Qy 841 GKDGSIYTSQIDKILQDKKDGTVITTSQELKKLADENKSKLSASDIASSINKLVGSMAL 900
 Db 841 GKDGSIYTSQIDKILQDKKDGTVITTSQELKKLADENKSKLSASDIASSINKLVGSMAL 900
 Qy 901 FGTANSVSSNALQPIPTQOGIILAPSV 927.
 Db 901 FGTANSVSSNALQPIPTQOGIILAPSV 927
 RESULT 2
 Q46716 PRELIMINARY; PRT; 998 AA.
 ID Q46716
 AC Q46716;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)

GN EHEC-HLYA.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95172699; PubMed=7868227;
 RA Schmidt H., Beutin L., Karch H.;
 RT "Molecular analysis of the plasmid-encoded hemolysin of Escherichia
 coli 0157:H7 strain Ed1933.";
 RL Infect. Immun. 63:1055-1061(1995).
 DR EMBL; X79839; CAA56234.1; -;
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR InterPro; IPR003995; RtxA.
 DR Pfam; PF00353; hemolysinCabin; 6.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PR00313; CABNDNGRPT.
 DR PRINTS; PR01488; RTXTOXINA.
 SQ SEQUENCE 998 AA; 107058 MW; 0D3BE108C309BBB3 CRC64;
 Query Match 1.5%; Score 14; DB 2; Length 998;
 Best Local Similarity 100.0%; Pred.No. 0.00015;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 379 ISGILEASKQAMFE 392
 DB 392 ISGILEASKQAMFE 405
 ID Q47461 PRELIMINARY; PRT; 998 AA.
 AC Q47461;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE EHEC-HLYA protein.
 GN EHEC-HLYA.
 OS Escherichia coli.
 OG Plasmid p0157.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=EDL 933;
 RX MEDLINE=95172699; PubMed=7868227;
 RA Schmidt H., Beutin L., Karch H.;
 RT "Molecular analysis of the plasmid-encoded hemolysin of Escherichia
 coli 0157:H7 strain Ed1933.";
 RL Infect. Immun. 63:1055-1061(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=EDL 933;
 RX MEDLINE=97090409; PubMed=8936317;
 RA Schmidt H., Kernbach C., Karch H.;
 RT "Analysis of the EHEC hly operon and its location in the physical map
 of the large plasmid of enterohemorrhagic Escherichia coli 0157:H7.";
 RL Microbiology 142:907-914(1996).
 DR EMBL; X86087; CAA60042.1; -;
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR InterPro; IPR003995; RtxA.
 DR Pfam; PF00353; hemolysinCabin; 6.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PR00313; CABNDNGRPT.
 DR PRINTS; PR01488; RTXTOXINA.
 KW Plasmid.
 SQ SEQUENCE 998 AA; 107032 MW; 4D8BE108C309BF7E CRC64;
 Query Match 1.5%; Score 14; DB 2; Length 998;
 Best Local Similarity 100.0%; Pred.No. 0.00015;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 379 ISGILEASKQAMFE 392
 DB 392 ISGILEASKQAMFE 405
 ID Q9LC58 PRELIMINARY; PRT; 998 AA.
 AC Q9LC58;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hemolysin A.
 GN EHEC-HLYA.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Taneike I., Wakisaka-Saito N., Harada Y., Zhang H.-M., Yamamoto T.;
 RT "The enterohemorrhagic Escherichia coli (EHEC)-hemolysin genes of a
 Shiga toxin 1 (Stx1)- and Stx2-producing, serotype O128 Escherichia
 coli strain with a greatest hemolytic activity.";
 RL Acta Med. Biol. (Niigata) 0:0-0(2000).
 DR EMBL; AB032930; BAA93708.1; -;
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR InterPro; IPR003995; RtxA.
 DR Pfam; PF00353; hemolysinCabin; 6.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PR00313; CABNDNGRPT.
 DR PRINTS; PR01488; RTXTOXINA.
 DR PROSITE; PS00330; HEMOLYSIN-CALCIUM; 1.
 SQ SEQUENCE 998 AA; 107196 MW; CD7A88E9BD862DB6 CRC64;
 Query Match 1.5%; Score 14; DB 2; Length 998;
 Best Local Similarity 100.0%; Pred.No. 0.00015;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 379 ISGILEASKQAMFE 392
 DB 392 ISGILEASKQAMFE 405
 ID Q8GA40 PRELIMINARY; PRT; 1024 AA.
 AC Q8GA40;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hemolysin A.
 GN HLYA.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=536;
 RX MEDLINE=22267134; PubMed=12379716;
 RA Dobrindt U., Blum-Oehler G., Nagy G., Schneider G., Johann A.,
 RA Gottschalk G., Hacker J.;
 RT "Genetic structure and distribution of four pathogenicity islands (PAI
 1536 to PAI IV536) of uropathogenic Escherichia coli strain 536.";
 RL Infect. Immun. 70:6365-6372(2002).
 DR EMBL; AJ488511; CAD33759.1; -;
 SQ SEQUENCE 1024 AA; 110314 MW; 9771FOA61ECF6EEB CRC64;
 Query Match 1.5%; Score 14; DB 2; Length 1024;
 Best Local Similarity 100.0%; Pred.No. 0.00015;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 379 ISGILEASKQAMPE 392
Db 406 ISGILEASKQAMPE 419
|||||
RESULT 9
Q8G9Z4 PRELIMINARY; PRT; 1024 AA.
AC Q8G9Z4
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE HlyA protein.
GN HlyA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=536; PubMed=12379716;
RX MEDLINE=22267134; PubMed=12379716;
RA Dobrindt U., Blum-Oehler G., Nagy G., Schneider G., Johann A.,
RA Gottschalk G., Hacker J.;
RT "Genetic structure and distribution of four pathogenicity islands (PAI
RT 1536 to PAI IV536) of uropathogenic Escherichia coli strain 536.";
RL Infect. Immun. 70:6365-6372(2002).
DR EMBL; AJ494981; CAD42039.1; -.
SQ SEQUENCE 1024 AA; 110365 MW; 3F3C284A45FEFE3 CRC64;

Query Match 1.5%; Score 14; DB 2; Length 1024;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 379 ISGILEASKQAMPE 392
Db 406 ISGILEASKQAMPE 419
|||||
RESULT 10
Q8FE01 PRELIMINARY; PRT; 1024 AA.
AC Q8FE01
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hemolysin A.
GN HlyA OR C3570.
OS Escherichia coli O6
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Rosesch P.,
RA Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AE016766; AA082018.1; -.
KW Complete proteome.
SQ SEQUENCE 1024 AA; 110246 MW; 9DEBA50C23B5A49C CRC64;

Query Match 1.5%; Score 14; DB 16; Length 1024;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 379 ISGILEASKQAMPE 392

```

```

Db 406 ISGILEASKQAMPE 419
|||||
RESULT 11
Q9RCG8 PRELIMINARY; PRT; 1049 AA.
AC Q9RCG8
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE PAXA.
GN PAXA.
OS Pasteurella aerogenes.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=749;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=field strain JF1319;
RA Heyberger-Meyer B., Frey J., Nicolet J., Kuhnert P.;
RT "Identification and characterization of a new RTX toxin operon (pax)
RT in Pasteurella aerogenes.";
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U66588; AAF15370.1; -.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCbind; 6.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 1.
SQ SEQUENCE 1049 AA; 112308 MW; 9FA5070E48CC3127 CRC64;

Query Match 1.5%; Score 14; DB 2; Length 1049;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 KVAAGFELSNOVIG 254
Db 275 KVAAGFELSNOVIG 288
|||||
RESULT 12
Q93NP0 PRELIMINARY; PRT; 1052 AA.
AC Q93NP0
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RTX-toxin IIIA.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=715;
RN [1]
RP SEQUENCE FROM N.A.
RA Shin S.J., Park J.Y., Choi I.S., Shin N.R., Yoo H.S.;
RT "Cloning and Sequencing of Apx IIIA from Actinobacillus
RT pleuropneumoniae.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF363363; AAK50053.1; -.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCbind; 6.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 1.
SQ SEQUENCE 1052 AA; 112862 MW; CA0E160F02101F42 CRC64;

Query Match 1.5%; Score 14; DB 2; Length 1052;
Best Local Similarity 100.0%; Pred. No. 0.00015;

```

```
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 KVAAGFELSNOVIG 254
Db 275 KVAAGFELSNOVIG 288
|||||

RESULT 13
Q8KWZ9 PRELIMINARY; PRT; 987 AA.
AC Q8KWZ9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE AQA.
GN AQA.
OS Actinobacillus cf. equuli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=182091;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCUG 19799;
RX MEDLINE=22031201; PubMed=12034544;
RA Berthoud H., Frey J., Kuhnert P.;
RT "Characterization of Aqx and its operon: the hemolytic RTX determinant
of Actinobacillus equuli."
RL Vet. Microbiol. 87:159-174 (2002).
DR EMBL; AF381184; AAM45566.1; -.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR Pfam; PF00353; hemolysinCabin; 1.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ SEQUENCE 987 AA; 105662 MW; CC34CD62BF9B745D CRC64;

Query Match 1.3%; Score 12; DB 2; Length 987;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 243 AAGFELSNOVIG 254
Db 277 AAGFELSNOVIG 288
|||||

RESULT 14
Q8KWZ6 PRELIMINARY; PRT; 987 AA.
AC Q8KWZ6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE AQA.
GN AQA.
OS Actinobacillus equuli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=90-570;
RX MEDLINE=22031201; PubMed=12034544;
RA Berthoud H., Frey J., Kuhnert P.;
RT "Characterization of Aqx and its operon: the hemolytic RTX determinant
of Actinobacillus equuli."
RL Vet. Microbiol. 87:159-174 (2002).
DR EMBL; AF381185; AAM45569.1; -.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR Pfam; PF00353; hemolysinCabin; 3.
DR Pfam; PF02382; RTX; 1.
```

```
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
SQ SEQUENCE 987 AA; 105871 MW; 35440030AB522E53 CRC64;

Query Match 1.3%; Score 12; DB 2; Length 987;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 243 AAGFELSNOVIG 254
Db 277 AAGFELSNOVIG 288
|||||

RESULT 15
Q93NP1 PRELIMINARY; PRT; 956 AA.
ID Q93NP1
AC Q93NP1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RTX toxin IIA.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=715;
RN [1]
RP SEQUENCE FROM N.A.
RA Shin S.J., Park J.Y., Choi I.S., Shin N.R., Yoo H.S.;
RT "Cloning and Sequencing of Apx IIA from Actinobacillus
pleuropneumoniae."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF363362; AAK50052.1; -.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ SEQUENCE 956 AA; 102495 MW; 9F8EB14EC76628D5 CRC64;

Query Match 1.2%; Score 11; DB 2; Length 956;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 369 ALLVAGVTGLI 379
Db 392 ALLVAGVTGLI 402
|||||

RESULT 16
Q93NP2 PRELIMINARY; PRT; 1022 AA.
ID Q93NP2
AC Q93NP2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RTX toxin IA.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=715;
RN [1]
RP SEQUENCE FROM N.A.
RA Shin S.J., Park J.Y., Choi I.S., Shin N.R., Yoo H.S.;
RT "Cloning and Sequencing of Apx IA from Actinobacillus
pleuropneumoniae."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF363361; AAK50051.1; -.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 6.
DR Pfam; PF02382; RTX; 1.
```

```

DR PRINTS: PR00113; CABNDNGRPT.
DR PRINTS: PR01488; RPYTOXINA.
SQ SEQUENCE 1022 AA; 110439 MW; A69E486D335467E3 CRC64;

Query Match 1.2%; Score 11; DB 2; Length 1022;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 266 AQRVAGLSTT 276
Db 289 AQRVAGLSTT 299

RESULT 17
Q8MR60
ID Q8MR60 PRELIMINARY; PRT; 593 AA.
AC Q8MR60;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE GH12004p.
GN CG5833 OR CGI1312.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY122108; AA52620.1; -.
DR Flybase; FBgn0032180; CG5833.
SQ SEQUENCE 593 AA; 61037 MW; 5AA228848C459627 CRC64;

Query Match 1.1%; Score 10; DB 5; Length 593;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 353 AGVSAAAVGS 362
Db 275 AGVSAAAVGS 284

RESULT 18
Q9ZFH0
ID Q9ZFH0 PRELIMINARY; PRT; 874 AA.
AC Q9ZFH0;
DT 01-MAY-1999 (TrEMBLrel. 10; Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mannuronan C-5-epimerase AlgE6.
GN ALGE6.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E;
RX MEDLINE=90084941; PubMed=9864314;
RA Svanem B.I.G., Skjaak-Braek G., Ertesvaag H., Valla S.;
RT "Cloning and expression of three new Azotobacter vinelandii genes
RT closely related to a previously described gene family encoding
RT mannuronan C-5-epimerases.";
RL J. Bacteriol. 181:68-77(1999).
DR EMBL; AF099799; AA004920.1; -.

```

```

DR InterPro; IPR006633; CASH.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR006626; PpHL.
DR Pfam; PF00353; hemolysinCbind; 8.
DR SMART; SM00722; CASH; 2.
DR SMART; SM00710; PpHL; 8.
SQ SEQUENCE 874 AA; 90159 MW; 22487F3A445427F1 CRC64;

Query Match 1.1%; Score 10; DB 2; Length 874;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 714 GDDLDDGGAG 723
Db 413 GDDLDDGGAG 422

RESULT 19
Q8RMZ8
ID Q8RMZ8 PRELIMINARY; PRT; 889 AA.
AC Q8RMZ8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 91.4 kDa protein.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RA Svanem B.I., Valla S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF486646; AAL93195.1; -.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR Pfam; PF00353; hemolysinCbind; 8.
KW Hypothetical protein.
SQ SEQUENCE 889 AA; 91391 MW; 6945D3971E5410D5 CRC64;

Query Match 1.1%; Score 10; DB 2; Length 889;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 714 GDDLDDGGAG 723
Db 608 GDDLDDGGAG 617

RESULT 20
Q9EV24
ID Q9EV24 PRELIMINARY; PRT; 946 AA.
AC Q9EV24;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LkTA.
OS Mannheimia glucosida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=85401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH574;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT haemolytica.";
RL J. Bacteriol. 183:1394-1404(2001).
DR EMBL; AF314521; AAG40305.1; -.
DR InterPro; IPR001753; EnCoA_hydrase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.

```

```

DR InterPro: IPR003995; RtxA.
DR Pfam: PF00353; hemolysinCabin; 5.
DR Pfam: PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ SEQUENCE 946 AA; 101480 MW; 25C077858BDC76C4 CRC64;

Query Match 1.1%; Score 10; DB 2; Length 946;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 276 LAQRVAAGLS 285

RESULT 21
Q9EV31 PRELIMINARY; PRT; 953 AA.
AC Q9EV31;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH588;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
haemolytica.";
RL J. Bacteriol. 183:1394-1404 (2001).
DR EMBL; AF314510; AAC0294.1; -.
DR InterPro: IPR001753; EnCoA_hydrase.
DR InterPro: IPR001343; Hemlysn_Ca_bind.
DR InterPro: IPR003995; RtxA.
DR Pfam: PF00353; hemolysinCabin; 5.
DR Pfam: PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102160 MW; A189BF80754A7907 CRC64;

Query Match 1.1%; Score 10; DB 2; Length 953;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 283 LAQRVAAGLS 292

RESULT 22
Q9ETX2 PRELIMINARY; PRT; 953 AA.
AC Q9ETX2;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Mannheimia glucosida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=85401;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=PH498, and PH344;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence diversity and molecular evolution of the leukotoxin (lktA)
gene in bovine and ovine strains of Mannheimia (Pasteurella)
haemolytica.";
RL J. Bacteriol. 0:0-0 (2001).
DR EMBL; AF314518; AAC0302.1; -.
DR EMBL; AF314517; AAC0301.1; -.
DR InterPro: IPR001753; EnCoA_hydrase.
DR InterPro: IPR001343; Hemlysn_Ca_bind.
DR InterPro: IPR003995; RtxA.
DR Pfam: PF00353; hemolysinCabin; 5.
DR Pfam: PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102135 MW; 70DB354157F5881E CRC64;

Query Match 1.1%; Score 10; DB 2; Length 953;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 283 LAQRVAAGLS 292

RESULT 23
Q9EV27 PRELIMINARY; PRT; 953 AA.
AC Q9EV27;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH392;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
haemolytica.";
RL J. Bacteriol. 183:1394-1404 (2001).
DR EMBL; AF314516; AAC0300.1; -.
DR InterPro: IPR001753; EnCoA_hydrase.
DR InterPro: IPR001343; Hemlysn_Ca_bind.
DR InterPro: IPR003995; RtxA.
DR Pfam: PF00353; hemolysinCabin; 5.
DR Pfam: PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102230 MW; 2B686808EB370090 CRC64;

Query Match 1.1%; Score 10; DB 2; Length 953;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 283 LAQRVAAGLS 292

RESULT 24
Q9EV29 PRELIMINARY; PRT; 953 AA.
ID Q9EV29

```

AC Q9EV29;
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Leukotoxin.
 GN LKTA.
 OS Pasteurella haemolytica.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Mannheimia.
 OX NCBI_TaxID=75985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PH278;
 RX MEDLINE=21101823; PubMed=11157953;
 RA Davies R.L., Whittam T.S., Selander R.K.;
 RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
 RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
 RT haemolytica.";
 RL J. Bacteriol. 183:1394-1404(2001).
 DR EMBL; AF314514; AAG40298.1; -;
 DR InterPro; IPR001753; EnCoA_hydrase.
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR InterPro; IPR003995; RtxA.
 DR Pfam; PF00353; hemolysinCabin; 5.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PR00313; CABNDNGRPT.
 DR PRINTS; PR01488; RTXTOXINA.
 DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
 SQ SEQUENCE 953 AA; 102147 MW; 11600FDA7849A1CA CRC64;

 Query Match 1.1%; Score 10; DB 2; Length 953;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 265 LAQRVAAGLS 274
 |||||
 DB 283 LAQRVAAGLS 292

 RESULT 25
 Q9EV25 PRELIMINARY; PRT; 953 AA.
 AC Q9EV25;
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Leukotoxin.
 GN LKTA.
 OS Mannheimia glucosida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Mannheimia.
 OX NCBI_TaxID=85401;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PH496;
 RX MEDLINE=21101823; PubMed=11157953;
 RA Davies R.L., Whittam T.S., Selander R.K.;
 RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
 RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
 RT haemolytica.";
 RL J. Bacteriol. 183:1394-1404(2001).
 DR EMBL; AF314520; AAG40304.1; -;
 DR InterPro; IPR001753; EnCoA_hydrase.
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR Pfam; PF00353; hemolysinCabin; 5.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PR00313; CABNDNGRPT.
 DR PRINTS; PR01488; RTXTOXINA.
 DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
 SQ SEQUENCE 953 AA; 102138 MW; C41D9EBC1D799951 CRC64;

 Query Match 1.1%; Score 10; DB 2; Length 953;

Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 265 LAQRVAAGLS 274
 |||||
 DB 283 LAQRVAAGLS 292

 RESULT 26
 Q9ETG5 PRELIMINARY; PRT; 953 AA.
 AC Q9ETG5;
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Leukotoxin.
 GN LKTA.
 OS Pasteurella haemolytica.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Mannheimia.
 OX NCBI_TaxID=75985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PH284, PH388, and PH8;
 RA Davies R.L., Whittam T.S., Selander R.K.;
 RT "Sequence diversity and molecular evolution of the leukotoxin (lktA)
 RT gene in bovine and ovine strains of Mannheimia (Pasteurella)
 RT haemolytica.";
 RL J. Bacteriol. 0:0-0(2001).
 DR EMBL; AF314507; AAG40291.1; -;
 DR EMBL; AF314504; AAG40288.1; -;
 DR EMBL; AF314506; AAG40290.1; -;
 DR InterPro; IPR001753; EnCoA_hydrase.
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR InterPro; IPR003995; RtxA.
 DR Pfam; PF00353; hemolysinCabin; 5.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PR00313; CABNDNGRPT.
 DR PRINTS; PR01488; RTXTOXINA.
 DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
 SQ SEQUENCE 953 AA; 101997 MW; D593D6A577C3ADE9 CRC64;

 Query Match 1.1%; Score 10; DB 2; Length 953;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 265 LAQRVAAGLS 274
 |||||
 DB 283 LAQRVAAGLS 292

 RESULT 27
 Q9EV23 PRELIMINARY; PRT; 953 AA.
 AC Q9EV23;
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Leukotoxin.
 GN LKTA.
 OS Mannheimia glucosida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Mannheimia.
 OX NCBI_TaxID=85401;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PH290;
 RX MEDLINE=21101823; PubMed=11157953;
 RA Davies R.L., Whittam T.S., Selander R.K.;
 RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
 RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
 RT haemolytica.";
 RL J. Bacteriol. 183:1394-1404(2001).


```

DR EMBL; AF314522; AAG40306.1; -
DR InterPro; IPR001753; EnCoA_hydratase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102150 MW; D99C36DA595B1624 CRC64;

Query Match 1.1%; Score 10; DB 2; Length 953;
Best Local Similarity 100.0%; Pred.No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 283 LAQRVAAGLS 292
|||||

RESULT 28
Q9EV33 PRELIMINARY; PRT; 953 AA.
AC Q9EV33;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PH66;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT haemolytica.";
RL J. Bacteriol. 183:1394-1404(2001).
DR EMBL; AF314508; AAG40292.1; -
DR InterPro; IPR001753; EnCoA_hydratase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102132 MW; 4138AB5FAE2843B3 CRC64;

Query Match 1.1%; Score 10; DB 2; Length 953;
Best Local Similarity 100.0%; Pred.No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 283 LAQRVAAGLS 292
|||||

RESULT 29
Q9EV34 PRELIMINARY; PRT; 953 AA.
AC Q9EV34;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

```

```

OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PH56;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT haemolytica.";
RL J. Bacteriol. 183:1394-1404(2001).
DR EMBL; AF314505; AAG40289.1; -
DR InterPro; IPR001753; EnCoA_hydratase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102010 MW; DE48B28EE0EB09FB CRC64;

Query Match 1.1%; Score 10; DB 2; Length 953;
Best Local Similarity 100.0%; Pred.No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 283 LAQRVAAGLS 292
|||||

RESULT 30
Q9EV32 PRELIMINARY; PRT; 953 AA.
AC Q9EV32;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PH706;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT haemolytica.";
RL J. Bacteriol. 183:1394-1404(2001).
DR EMBL; AF314509; AAG40293.1; -
DR InterPro; IPR001753; EnCoA_hydratase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102078 MW; EF425243C8741EE4 CRC64;

Query Match 1.1%; Score 10; DB 2; Length 953;
Best Local Similarity 100.0%; Pred.No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 283 LAQRVAAGLS 292
|||||

```

```

RESULT 31
Q9EV30 PRELIMINARY; PRT; 953 AA.
AC Q9EV30;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH196;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT haemolytica."
RL J. Bacteriol. 183:1394-1404 (2001).
DR EMBL; AF314512; AAC40296.1; -.
DR InterPro; IPR001753; EnCoA_hydrase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102043 MW; 4E8F11490479A69A CRC64;

Query Match 1.1%; Score 10; DB 2; Length 953;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQRVAAGLS 274
DB 283 LAQRVAAGLS 292
|||||

RESULT 32
Q9EV28 PRELIMINARY; PRT; 953 AA.
AC Q9EV28;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Leukotoxin (lktA).
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH292, and PH296;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT haemolytica."
RL J. Bacteriol. 183:1394-1404 (2001).
DR EMBL; AF314515; AAG40299.1; -.
DR InterPro; IPR001753; EnCoA_hydrase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102043 MW; 4E8F11490479A69A CRC64;

Query Match 1.1%; Score 10; DB 2; Length 953;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQRVAAGLS 274
DB 283 LAQRVAAGLS 292
|||||

RESULT 33
Q9EV26 PRELIMINARY; PRT; 953 AA.
AC Q9EV26;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Mannheimia glucosida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=85401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH240;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT haemolytica."
RL J. Bacteriol. 183:1394-1404 (2001).
DR EMBL; AF314519; AAG40303.1; -.
DR InterPro; IPR001753; EnCoA_hydrase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102236 MW; AC5874B3B75D1C98 CRC64;

Query Match 1.1%; Score 10; DB 2; Length 953;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQRVAAGLS 274
DB 283 LAQRVAAGLS 292
|||||

RESULT 34
Q9EUE1 PRELIMINARY; PRT; 954 AA.
AC Q9EUE1;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

```

```

DR EMBL; AF414141; AAL13281.1; -.
DR InterPro; IPR001753; EnCoA_hydrase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102218 MW; 04AB1715B819E571 CRC64;

Query Match 1.1%; Score 10; DB 2; Length 953;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQRVAAGLS 274
DB 283 LAQRVAAGLS 292
|||||

RESULT 35
Q9EUE1 PRELIMINARY; PRT; 954 AA.
AC Q9EUE1;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

```

```

OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH202, and PH494;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence diversity and molecular evolution of the leukotoxin (lktA)
RT gene in bovine and ovine strains of Mannheimia (Pasteurella)
RT haemolytica.";
RL J. Bacteriol. 0:0-0(2001).
DR EMBL; AF314513; AAG40297.1; -.
DR EMBL; AF314511; AAG40295.1; -.
DR InterPro; IPR001753; EnCoA_hydratase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 954 AA; 102410 MW; 02A60E52411711A2 CRC64;

Query Match 1.1%; Score 10; DB 2; Length 954;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 283 LAQRVAAGLS 292
|||||

RESULT 35
Q9EUD4 PRELIMINARY; PRT; 955 AA.
ID Q9EUD4;
AC Q9EUD4;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence diversity and molecular evolution of the leukotoxin (lktA)
RT gene in bovine and ovine strains of Mannheimia (Pasteurella)
RT haemolytica.";
RL J. Bacteriol. 0:0-0(2001).
DR EMBL; AF314525; AAG40309.1; -.
DR EMBL; AF314523; AAG40307.1; -.
DR EMBL; AF314524; AAG40308.1; -.
DR InterPro; IPR001753; EnCoA_hydratase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 955 AA; 102127 MW; 074B2E4ADFF57AA CRC64;

Query Match 1.1%; Score 10; DB 2; Length 955;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 285 LAQRVAAGLS 294
|||||

Pasteurella trehalosi.
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Pasteurella.
NCBI_TaxID=47735;
[1]
SEQUENCE FROM N.A.
STRAIN=PH254, and PH252;
Davies R.L., Whittam T.S., Selander R.K.;
"Sequence diversity and molecular evolution of the leukotoxin (lktA)
gene in bovine and ovine strains of Mannheimia (Pasteurella)
haemolytica.";
J. Bacteriol. 0:0-0(2001).
EMBL; AF314525; AAG40309.1; -.
EMBL; AF314523; AAG40307.1; -.
EMBL; AF314524; AAG40308.1; -.
InterPro; IPR001753; EnCoA_hydratase.
InterPro; IPR001343; Hemlysn_Ca_bind.
InterPro; IPR003995; RtxA.
Pfam; PF00353; hemolysinCabin; 5.
Pfam; PF02382; RTX; 1.
PRINTS; PR00313; CABNDNGRPT.
PRINTS; PR01488; RTXTOXINA.
PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SEQUENCE 955 AA; 102127 MW; 074B2E4ADFF57AA CRC64;

Query Match 1.1%; Score 10; DB 2; Length 955;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 285 LAQRVAAGLS 294
|||||

```

```

RESULT 36
Q9EV22 PRELIMINARY; PRT; 955 AA.
ID Q9EV22;
AC Q9EV22;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence diversity and molecular evolution of the leukotoxin (lktA)
RT gene in bovine and ovine strains of Mannheimia (Pasteurella)
RT haemolytica.";
RL J. Bacteriol. 183:1394-1404(2001).
DR EMBL; AF314526; AAG40310.1; -.
DR InterPro; IPR001753; EnCoA_hydratase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 955 AA; 102069 MW; FDC2E99B8DC394F CRC64;

Query Match 1.1%; Score 10; DB 2; Length 955;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 285 LAQRVAAGLS 294
|||||

Pasteurella trehalosi.
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Pasteurella.
NCBI_TaxID=47735;
[1]
SEQUENCE FROM N.A.
STRAIN=PH68;
MEDLINE=21101823; PubMed=11157953;
Davies R.L., Whittam T.S., Selander R.K.;
"Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
haemolytica.";
J. Bacteriol. 183:1394-1404(2001).
DR EMBL; AF314526; AAG40310.1; -.
DR InterPro; IPR001753; EnCoA_hydratase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 955 AA; 102069 MW; FDC2E99B8DC394F CRC64;

Query Match 1.1%; Score 10; DB 2; Length 955;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 285 LAQRVAAGLS 294
|||||

Pasteurella trehalosi.
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Pasteurella.
NCBI_TaxID=47735;
[1]
SEQUENCE FROM N.A.
STRAIN=E;
MEDLINE=94245609; PubMed=8188585;
Ertesvag H., Doseth B., Larsen B., Skjak-Braek G., Valla S.;
"Cloning and expression of an Azotobacter vinelandii mannuronan C-5-
epimerase gene";
J. Bacteriol. 176:2846-2853(1994).
[2]
SEQUENCE FROM N.A.
STRAIN=E;
MEDLINE=96065700; PubMed=7476166;
Ertesvag H., Hoidal H.K., Hale I.K., Rian A., Doseth B., Valla S.;
"A family of modular type mannuronan C-5-epimerase genes controls
arginate structure in Azotobacter vinelandii.";
Mol. Microbiol. 16:719-731(1995).
EMBL; L39096; AAA87311.1; -.

```

```

DR InterPro; IPR006633; CASH.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR DR Pfam; PF00353; hemolysinCbind; 12.
DR DR SMART; SM00722; CASH; 4.
DR DR SMART; SM00710; PBH1; 15.
SQ SEQUENCE 1403 AA; 147168 MW; 4B843AB0A366A95C CRC64;

Query Match 1.1%; Score 10; DB 2; Length 1403;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 714 GDLLDGGAG 723
Db 1257 GDLLDGGAG 1266
|||||

RESULT 38
Q9VL42 PRELIMINARY; PRT; 1424 AA.
AC Q9VL42;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG31714 protein.
GN CS5833 OR CG3132.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J., Blazek R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Bao A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolehakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Flier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert B.C., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195(2000).

```

```

[2] SEQUENCE FROM N.A.
RN RP Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferrieria S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C.A., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RN RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
RN RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS003627; AAF52856.2; -.
DR FlyBase; FBgn0032180; CG5833.
DR InterPro; IPR001879; hormn_receptor.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF02793; HRN; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50227; G_PROTEIN_RECEPT_F2_3; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
SQ SEQUENCE 1424 AA; 152768 MW; E8E39E276FD7414E CRC64;

Query Match 1.1%; Score 10; DB 5; Length 1424;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 353 AGVSAAGVS 362
Db 1106 AGVSAAGVS 1115
|||||

RESULT 39
Q8YKJ3 PRELIMINARY; PRT; 4936 AA.
AC Q8YKJ3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein Alr7304.
GN ALR7304.
OS Anabaena sp. (strain PCC 7120).
OC Plasmid pCC7120alpha.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,

```

RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL; AP003600; BAB78388.1; -.
 DR InterPro; IPR006644; Cadg.
 DR InterPro; IPR003644; Calx_beta.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF03160; Calx_beta; 5.
 DR Pfam; PF00353; hemolysinCbind; 25.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00313; CAENDNGRPT.
 DR PRINTS; PR00723; SUBTILISIN.
 DR SMART; SM00736; CADG; 2.
 DR SMART; SM00237; Calx_beta; 6.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 2.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 KW Plasmid: Hypothetical protein; Complete proteome.
 SQ SEQUENCE 4936 AA; 519407 MW; F5A8323CB29C828 CRC64;

Query Match 1.1%; Score 10; DB 16; Length 4936;
 Best Local Similarity 100.0%; Pred. No. 8;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 752 DVLNGGAGND 761
 |||||
 Db 4296 DVLNGGAGND 4305

RESULT 40
 Q9RDA8 PRELIMINARY; PRT; 259 AA.
 AC Q9RDA8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein SCO2843.
 GN SCO2843 OR SCE20.17C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,

RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL; AL939114; CAB65573.1; -.
 DR InterPro; IPR006357; HAD_SF_IIA.
 DR InterPro; IPR005834; Hydrolase.
 DR Pfam; PF00702; Hydrolase; 1.
 DR TIGRfams; TIGR01460; HAD-SF-IIIa; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 259 AA; 28059 MW; 997D2583D4F7835A CRC64;

Query Match 1.0%; Score 9; DB 16; Length 259;
 Best Local Similarity 100.0%; Pred. No. 5.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 276 TGAVAAALIT 284
 |||||
 Db 164 TGAVAAALIT 172

Search completed: February 17, 2004, 10:21:24
 Job time : 85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:09:49 ; Search time 21 Seconds
(without alignments)
1867.722 Million cell updates/sec

Title: US-10-069-799-5
Perfect score: 4647
Sequence: 1 MSINIVKSNQAGLNSTKS.....SSNALQPIPTQTOGILAPSV 927

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------------|
| 1 | 2334 | 50.2 | 1098 | 1 | US-07-777-715-7 |
| 2 | 2334 | 50.2 | 1098 | 1 | Sequence 7, Appli |
| 3 | 2334 | 50.2 | 1098 | 3 | Sequence 2, Appli |
| 4 | 2366.5 | 50.1 | 924 | 3 | Sequence 8, Appli |
| 5 | 2326.5 | 50.1 | 926 | 1 | Sequence 2, Appli |
| 6 | 2326.5 | 50.1 | 926 | 1 | Sequence 2, Appli |
| 7 | 2326.5 | 50.1 | 926 | 2 | Sequence 6, Appli |
| 8 | 2326.5 | 50.1 | 926 | 2 | Sequence 6, Appli |
| 9 | 2326.5 | 50.1 | 926 | 2 | Sequence 6, Appli |
| 10 | 2326.5 | 50.1 | 926 | 2 | Sequence 2, Appli |
| 11 | 2326.5 | 50.1 | 926 | 3 | Sequence 6, Appli |
| 12 | 2326.5 | 50.1 | 926 | 6 | Patent No. 5476657 |
| 13 | 2326.5 | 50.1 | 926 | 1 | Sequence 12, Appl |
| 14 | 2326.5 | 50.1 | 943 | 1 | Sequence 10, Appl |
| 15 | 2326.5 | 50.1 | 951 | 1 | Sequence 14, Appl |
| 16 | 2326.5 | 50.1 | 977 | 1 | Sequence 8, Appli |
| 17 | 2326.5 | 50.1 | 977 | 2 | Sequence 8, Appli |
| 18 | 2326.5 | 50.1 | 977 | 2 | Sequence 8, Appli |
| 19 | 2326.5 | 50.1 | 977 | 3 | Sequence 8, Appli |
| 20 | 2326.5 | 50.1 | 977 | 3 | Sequence 8, Appli |
| 21 | 2326.5 | 50.1 | 1069 | 1 | Sequence 9, Appli |
| 22 | 2326.5 | 50.1 | 1069 | 1 | Sequence 4, Appli |
| 23 | 2326.5 | 50.1 | 1069 | 3 | Sequence 4, Appli |
| 24 | 2326.5 | 50.1 | 1069 | 3 | Sequence 80, Appl |
| 25 | 2297.5 | 49.4 | 934 | 1 | Sequence 8, Appli |
| 26 | 2236.5 | 48.1 | 956 | 3 | Sequence 8, Appli |
| 27 | 2236.5 | 48.1 | 956 | 4 | Sequence 8, Appli |

| | | | | | | |
|----|--------|------|------|---|-------------------|-------------------|
| 28 | 1926 | 41.4 | 1053 | 4 | US-09-062-126-10 | Sequence 10, Appl |
| 29 | 1924 | 41.4 | 1049 | 3 | US-08-772-270A-11 | Sequence 11, Appl |
| 30 | 1924 | 41.4 | 1244 | 5 | PCT-US93-10500-2 | Sequence 2, Appli |
| 31 | 1763.5 | 37.9 | 1022 | 3 | US-08-772-270A-2 | Sequence 2, Appli |
| 32 | 1751.5 | 37.7 | 1013 | 4 | US-09-062-126-3 | Sequence 3, Appli |
| 33 | 1376.5 | 29.6 | 758 | 1 | US-08-258-188-2 | Sequence 2, Appli |
| 34 | 1376.5 | 29.6 | 758 | 1 | US-08-526-813-2 | Sequence 2, Appli |
| 35 | 1376.5 | 29.6 | 758 | 5 | PCT-US95-08554-2 | Sequence 2, Appli |
| 36 | 904.5 | 19.5 | 544 | 1 | US-08-387-156-10 | Sequence 10, Appl |
| 37 | 904.5 | 19.5 | 544 | 2 | US-08-694-865-10 | Sequence 10, Appl |
| 38 | 904.5 | 19.5 | 544 | 2 | US-08-878-748-10 | Sequence 10, Appl |
| 39 | 904.5 | 19.5 | 544 | 3 | US-09-124-491-10 | Sequence 10, Appl |
| 40 | 904.5 | 19.5 | 544 | 4 | US-09-383-912-10 | Sequence 16, Appl |
| 41 | 901 | 19.4 | 699 | 2 | US-08-694-865-16 | Sequence 16, Appl |
| 42 | 901 | 19.4 | 699 | 3 | US-09-124-491-16 | Sequence 16, Appl |
| 43 | 901 | 19.4 | 699 | 4 | US-09-383-912-16 | Sequence 16, Appl |
| 44 | 897.5 | 19.3 | 490 | 4 | US-09-252-149B-26 | Sequence 26, Appl |
| 45 | 896.5 | 19.3 | 1403 | 1 | US-07-908-253-3 | Sequence 3, Appli |

ALIGNMENTS

RESULT 1
US-07-777-715-7
; Sequence 7, Application US/07777715
; Patent No. 5273889
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew
; APPLICANT: Campos, Manuel
; APPLICANT: Hughes, Huw P.A.
; TITLE OF INVENTION: CYTOKINE-LEUKOTOXIN GENE FUSIONS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/777,715
; FILING DATE: 19911016
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Robins, Roberta L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 29310-2001320
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; TELEFAX: 415-327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-777-715-7

Query Match 50.2% Score 2334; DB 1; Length 1098;
Best Local Similarity 50.1% Pred. No. 5.2e-174;
Matches 465; Conservative 175; Mismatches 263; Indels 26; Gaps 13;
Qy 8 KSNITQAGLNKSGKLNLYLAIPKD--YDPKQGTGLNDFTKAADELGIARLAEPNHTET 65
Db 173 QSLTQAG-SSILKTGAKKIILVIPQNYQYDTQGNGLQDLVKADELGIETVQREERNNIAT 231

QY 66 AKKSDVTNQNFLSTOTGTIAISATKLEFKLOKSTNKLAKGLDSVENIDRKLKASNVLS 125
DB 232 AQTSLGTIQTATIGLGERGIVLAPQIDKLQK---TKAGQALGSAESIVQNKAKTVLS 288
QY 126 TLSSFLGTALAGIELDLSIKKGDADAPDALAKASIDLINIEITGNLSQSTQTIEAFSSQLAK 185
DB 289 GIQSLGSLVAGMDLDEAL-QNNSQHALAKAGLELTNSLIENIANSVKTLDEGEQISQ 347
QY 186 LGSTISQAKGFSNIGNKQNL-NFSKTNLGLIEITGLLSGISAGPALADKNAKSTGKVVAA 244
DB 348 FGSKLQNIKGLGTGLDKLKNIGLGDVLSGLSGATAALVLADKNAKSTAKKVG 407
QY 245 GFELSNQVGNVTKAISYVLAQRAAGLSTTGAVALITSSIMLAIPLAFMAADKFN 304
DB 408 GFELANQVGNVTKAISYVLAQRAAGLSTTGAVALITSSIMLAIPLAFMAADKFN 467
QY 305 HANALDEFKQPRKFGYDGDHLLAEYQRGVGTIEASLTITSTALGAVSAGVSAAGVSAV 364
DB 468 HAKSLESVAERFKLGYDGDHLLAEYQRGVGTIEASLTITSTALGAVSAGVSAAGVSAV 527
QY 365 GTPIALLVAGVTGLISGILEASKAMFESVANRQKLEWEKQNGGQNYFDKGYDSRYA 424
DB 528 ASPIALLVSGITGVISTILOYSKOAMFEHVANKIHKIVEKKNHGNFYFENGYDARYL 587
QY 425 AYLANNLKELSELNKELEAERVIATQORWNNIGELAGITKLERIKSGKAYADAFEDG 484
DB 588 ANLQDNMFKLLNLKELQAERVIATQOQWNNIGDLGAGISRLGEKVLGKAYVDAFEEG 647
QY 485 KKEVAGSNITLDAKTGIIDISNSNGKTKQALHFTSPLLTAGTESRRLTNGKYSYINKL 544
DB 648 KHIKADLVQDSANGIIDVNSGKAKTQHILFRPLTPGTEHREVRTGKYEYITKLN 707
QY 545 FGRVKMQVTGGEASSKLDPSKVIQV-----AETGTEIGLIVNAKAGNDDIFVGQ 597
DB 708 INRVDSWKITDGAASSTFDLNVVQRIIGELDNAGNVTKTKETKIIAKLGBGDDNVFVGS 767
QY 598 GGMNIDGCGDHRVFSKGGGNGTVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
DB 768 GTTEIDGGEYDRVHSYR-GNYGALITIDATKETEQQSYTVNRVFTGKALHEVSTHTAL 826
QY 658 VGRKRTTIQYRDYELRKVGYGQSYDNLKSVBEVIGSOFNDVFKGSKFNDIFHSGEQDL 717
DB 827 VGNREKEIYR-HSNQNHAGYTYKDTLKAVEEIIIGTSHNDIFKSKFNDAFNGDGYDT 885
QY 718 LDGAGDRLFGKGNDRLSGDEGDDLLDGGSGDDVNLGGAGNDVYIPRKGNDTLVDG 777
DB 886 IDGNDGNDRLFGKGGDDILDGGNGDDFDGGKNDLLHGGKDDIFVHRKGGDNDIIDS 945
QY 778 TGNDKLAADANISDIMERTKEGIIVKRNHSGSINIPRWY----ITSNLQYQSNKTD 833
DB 946 DGNKLSFSDSNLKDLPTEKVKHNLVI-TNSKKEKVTIQNFPREADFAKEVFNKATK-D 1003
QY 834 HKIEOLIGKGSYITSDQIDKIDKQDGTITTSOELKKLADENKSKLSASDIASSLNK 893
DB 1004 EKIEIILGGERIITSKVQDDLI--AKGNKTIQDELKVDVNYELLKHS-KVNTSLDX 1060
QY 894 LVGSMALFGTANSSNALQPIPTQPTQGI 922
DB 1061 LITSSVSAFTSSNDRSNVLVAPTSMLDQSL 1089

RESULT 2

US-08-170-126-2

Sequence 2, Application US/08170126

Patent No. 5594107

GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW

APPLICANT: CAMPOS, MANUEL

APPLICANT: HUGHES, HUW P.A.

TITLE OF INVENTION: CYTOKININ GENE FUSIONS AND USES

TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 6

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: REED & ROBINS
;; STREET: 635 BRYANT STREET
;; CITY: PALO ALTO
;; STATE: CALIFORNIA
;; COUNTRY: UNITED STATES OF AMERICA
;; ZIP: 94301
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/170,126
;; FILING DATE: 20-DEC-1993
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/777,715
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/571,301
;; ATTORNEY/AGENT INFORMATION:
;; NAME: ROBINS, ROBERTA L.
;; REGISTRATION NUMBER: 33,208
;; REFERENCE/DOCKET NUMBER: 9000-0013.21
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 617-8999
;; TELEFAX: (415) 327-3231
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1098 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-170-126-2

Query Match 50.2%; Score 2334; DB 1; Length 1098;

Best Local Similarity 50.1%; Pred. No. 5.2e-174;

Matches 465; Conservative 175; Mismatches 263; Indels 26; Gaps 13;

QY 8 KSNTOAGLNSTKSLKLYLAIPKD--YDPOKGGTLNDFIKAADELGIARLAEPNHTET 65
DB 173 QSLTQAG-SSLTKGAKIILYIPQNYQYDTEQNGLODLVKAABELGIEVQREERNIAT 231
QY 66 AKKSDVTNQNFLSTOTGTIAISATKLEFKLOKSTNKLAKGLDSVENIDRKLKASNVLS 125
DB 232 AQTSLGTIQTATIGLGERGIVLAPQIDKLQK---TKAGQALGSAESIVQNKAKTVLS 288
QY 126 TLSSFLGTALAGIELDLSIKKGDADAPDALAKASIDLINIEITGNLSQSTQTIEAFSSQLAK 185
DB 289 GIQSLGSLVAGMDLDEAL-QNNSQHALAKAGLELTNSLIENIANSVKTLDEGEQISQ 347
QY 186 LGSTISQAKGFSNIGNKQNL-NFSKTNLGLIEITGLLSGISAGPALADKNAKSTGKVVAA 244
DB 348 FGSKLQNIKGLGTGLDKLKNIGLGDVLSGLSGATAALVLADKNAKSTAKKVG 407
QY 245 GFELSNQVGNVTKAISYVLAQRAAGLSTTGAVALITSSIMLAIPLAFMAADKFN 304
DB 408 GFELANQVGNVTKAISYVLAQRAAGLSTTGAVALITSSIMLAIPLAFMAADKFN 467
QY 305 HANALDEFKQPRKFGYDGDHLLAEYQRGVGTIEASLTITSTALGAVSAGVSAAGVSAV 364
DB 468 HAKSLESVAERFKLGYDGDHLLAEYQRGVGTIEASLTITSTALGAVSAGVSAAGVSAV 527
QY 365 GTPIALLVAGVTGLISGILEASKAMFESVANRQKLEWEKQNGGQNYFDKGYDSRYA 424
DB 528 ASPIALLVSGITGVISTILOYSKOAMFEHVANKIHKIVEKKNHGNFYFENGYDARYL 587
QY 425 AYLANNLKELSELNKELEAERVIATQORWNNIGELAGITKLERIKSGKAYADAFEDG 484
DB 588 ANLQDNMFKLLNLKELQAERVIATQOQWNNIGDLGAGISRLGEKVLGKAYVDAFEEG 647
QY 485 KKEVAGSNITLDAKTGIIDISNSNGKTKQALHFTSPLLTAGTESRRLTNGKYSYINKL 544

Db 648 KHIKADKLVQDSANGIIVDSNSCKAKTQHILFRTPLLTPTGTEHRRVQTKYEYITKLN 707
Qy 545 FGRVKNQVTDGEASSKLDPSKVIQV-----AETEGTDEGLIIVNAKAGNDIDFVGQ 597
Db 708 INRYDSWKITDGAASSTFDLTNNVQIRIGIELDNAGNVTKTETKIIIAKLGEGDDNVFVGS 767
Qy 598 GKMNIDGGDHRVYFSKDGFGNITVDGTSATEAGSYTVNRKVARGDIIYHEVVKRQETK 657
Db 768 GTTEIDGEGYDRVHYSR-GNYGALTIDATKETEFGSYTVNRFVETGKALHEVTSHTAL 826
Qy 658 VGRKTEIQRDYBELRKVGVGYOSTNMLKSVVEVIGSQFNDVFKGSKFNDI FHSGEDDL 717
Db 827 VGNREEKIEYR-HSNNOHAGYTTKDTLKAVEEIIIGTSHNDIFPKSGKFNDAFNGDGVDT 885
Qy 718 LDGAGDRLFGGKGNDRSLSGEDGDLDDGSGDDVLNGAGNDVIYFRKGDDNDTLVDG 777
Db 886 IDGNDGNDRLFGGKGDIDLDGGNGDDFDGKGNDLLHGGKDDIFVHRKGDDNDIITDS 945
Qy 778 TGNDKLAFADANISDIMERKEGIIVKRNDHSGSINIPRWY-----ITSNLQYQSNKTD 833
Db 946 DGNDKLSFSDSNLKDITFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 1003
Qy 834 HKIBQLKDGSGYITSDQIDKILQDKDGTVTITSQELKKLADENKSKLSASDIASSLNK 893
Db 1004 EKIEEIIIGQGERITSKQVDDLI--AKNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 1060
Qy 894 LVGSMALFGTANSVSSNALQIPITQPTQGI 922
Db 1061 LISSVSAFTSSNDNRNVLVAPTSMLDQSL 1089

RESULT 3

US-08-954-418-2
; Sequence 2, Application US/08954418
; Patent No. 6096320
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW
; APPLICANT: CAMPOS, MANUEL
; APPLICANT: HUGHES, HUN P.A.
; TITLE OF INVENTION: CYTOKINE-CYTOTOXIN GENE FUSIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,418
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170,126
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/571,301
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9000-0013.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-954-418-2
Query Match 50.2%; Score 2334; DB 3; Length 1098;
Best Local Similarity 50.1%; Pred. No. 5,2e-174;
Matches 465; Conservative 175; Mismatches 263; Indels 26; Gaps 13;
Qy 8 KSNIOAGLNTSKGLKNLVIAPKD--YDPOKGGTINDFIKADELGIARLAEPNHET 65
Db 173 QSTQAG-SLKTGAKKILYIPQNYQYDEQNGLODLVKAABELGIEVQREERNIAT 231
Qy 66 AKKSVDVTNQFLSTGTGIAISATKLEKFLQKHSITNKLAKGLDSVENIDRKLKASNLS 125
Db 232 AQTSLGTIQTALGTERGIVLSAPQDKLQK--TKAQALGSAESIVQNKAKTVLS 288
Qy 126 TLSFSLGTALAGIELDSLKKGDAAPDALAKASIDLINEIIGNLSOSTQTIEAFSSOLAK 185
Db 289 GIQSILGSLVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDFEGEIQS 347
Qy 186 LGSTISOAKGFSNIGKNLQNL-NFSKTNLGLLEIITGLGISAGFALADKNASTGKKVAA 244
Db 348 FGSKLQNIKGLTGLDKLKNIGGLDKAGLGLDVISLGSATLALVLADKNASTAKKVG 407
Qy 245 GFELSNQVIGNVTKAISYVLAQVAAAGLSTTGAAVALITSSIMLAISPLAFNAADKFN 304
Db 408 GFELANQVGNITKAVSSYILAQRVAAGLSSTGTPVAALIASTVSLAISPLAFAGIADKFN 467
Qy 305 HANALDEFAPKPRFGYDGDHLLAEYQORGVTTEASLTITSTALGAVSAGVSAAVGSV 364
Db 468 HAKLSYEAERFKKLGVDGNLAEYQRGTTIDASVTAINTAALAAIAGGVSAAGASVI 527
Qy 365 GTPALLVAGVTGLISILKASKOMFESVANRLOKLEMEKONGCONFYDKGYSRYA 424
Db 528 ASPIALLVSGITGVISTILOYSKOAMFEHVANKHNKVEWKNHGNKNTFENGIDARYL 587
Qy 425 AYLANNLKFLSELNKELEAEERVIATQQRDNNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db 588 ANLQDNKFWLLNLKELQAEVIAITQOQWDNNIGDLAGISRLGEKVLGSKAYVDAFEEG 647
Qy 485 KYVEAGSNITLDAKTGIIIDISNGSKTQALHTSPPLTAGTESRELTNGKYSYINKLK 544
Db 648 KHIKADKLVQDSANGIIVDSNSCKAKTQHILFRTPLLTPTGTEHRRVQTKYEYITKLN 707
Qy 545 FGRVKNQVTDGEASSKLDPSKVIQV-----AETEGTDEGLIIVNAKAGNDIDFVGQ 597
Db 708 INRYDSWKITDGAASSTFDLTNNVQIRIGIELDNAGNVTKTETKIIIAKLGEGDDNVFVGS 767
Qy 598 GKMNIDGGDHRVYFSKDGFGNITVDGTSATEAGSYTVNRKVARGDIIYHEVVKRQETK 657
Db 768 GTTEIDGEGYDRVHYSR-GNYGALTIDATKETEFGSYTVNRFVETGKALHEVTSHTAL 826
Qy 658 VGRKTEIQRDYBELRKVGVGYOSTNMLKSVVEVIGSQFNDVFKGSKFNDI FHSGEDDL 717
Db 827 VGNREEKIEYR-HSNNOHAGYTTKDTLKAVEEIIIGTSHNDIFPKSGKFNDAFNGDGVDT 885
Qy 718 LDGAGDRLFGGKGNDRSLSGEDGDLDDGSGDDVLNGAGNDVIYFRKGDDNDTLVDG 777
Db 886 IDGNDGNDRLFGGKGDIDLDGGNGDDFDGKGNDLLHGGKDDIFVHRKGDDNDIITDS 945
Qy 778 TGNDKLAFADANISDIMERKEGIIVKRNDHSGSINIPRWY-----ITSNLQYQSNKTD 833
Db 946 DGNDKLSFSDSNLKDITFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 1003
Qy 834 HKIEQLKDGSGYITSDQIDKILQDKDGTVTITSQELKKLADENKSKLSASDIASSLNK 893
Db 1004 EKIEEIIIGQGERITSKQVDDLI--AKNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 1060
Qy 894 LVGSMALFGTANSVSSNALQIPITQPTQGI 922
Db 1061 LISSVSAFTSSNDNRNVLVAPTSMLDQSL 1089


```
RESULT 4
US-08-619-812-8
; Sequence 8, Application US/08619812
; Patent No. 6100066
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: THEISEN, MICHAEL
; APPLICANT: HARLAND, RICHARD J.
; APPLICANT: RIOUX, CLEMENT R.
; TITLE OF INVENTION: VACCINES FOR HAEMOPHILUS SOMNUS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,812
; FILING DATE: 15-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/038,719
; FILING DATE: 29-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9000-0019.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 924 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-619-812-8
Query Match 50.1%; Score 2326.5; DB 3; Length 924;
Best Local Similarity 50.0%; Pred. No. 1.5e-173;
Matches 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12;
QY 15 LNSTSGKLNLYLAIPKD--YDPQGGTLNDFIKAADELGIARLAEEPHNTHETAKKSVDT 72
DB 7 LSPFKTGAKKILYIPQYQYDTEQGNGLQDLVKAABELGIEVOREERNIATAQTSLGT 66
QY 73 VNQFLSLTQTGAISATKLEKFLQKHSNKLAKGLDSVENTDRKLGKASNVLSLSSFLG 132
DB 67 TQTAGLTERGIVLSAPQIDKLLQK--TKAGQALGSAESIVONANKAKTVLSGIQSLG 123
QY 133 TALAGIELDSLIKKGDAAPDALAKASIDLNIENLQSQTQTIERAFSSQLAKGLSTISQ 192
DB 124 SVLACMDLDEAL-QNNSNHQALAKAGLEITNSLIENIANSVKTLDFFGEQISQFGSKLQN 182
QY 193 AKGFNIGNKQLNL-NFSKTNLGLBIIITGLSGISAGFALADKNASTGKKVAAGFELSNQ 251
DB 183 INGLGTGLDKLKNIGLDKAGLGDVIGLISGATAALVLADKNASTAKVVGAGFELANQ 242
QY 252 VIGNVTKAISVYLAQRVAAGLSTTGAAVALTSSIMLAISPLAFMNAADKFNHANALDE 311
DB 243 VGNITKAVSSYLAQRVAAGLSTTGAAVALTSSIMLAISPLAFMNAADKFNHANALDE 302
QY 312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTITSTALGAVSAGVSAAGVAVGTPPIALL 371
DB 303 YAEFRFKLGDGDNLLAEYQRGVGTIEASLTITSTALGAVSAGVSAAGVAVGTPPIALL 362
```

```

372 VAGVTGLISGLEASKOAMRESVANRLQGLKLEWEKONGYFDKGYDSRYAAYLANNL 431
DB 363 VSGITGVISTILOYSKOAMFEHVANKIHNKIVVEKKNHGNKYFENGYDARYLANLQNM 422
QY 432 KFLSELNKELEAERVAITQORWNNITGELAGITKGLGERIKSGKAYADAFEDGKKVEAGS 491
DB 423 KFLNLNKELEAERVAITQOOWNNIGLAGISRLGKVLGKAYYDAFEGKHAKDK 482
QY 492 NITLDAKTIIDISNSNGKKTQALHFTSPILTAGTESRERLTNGKYSYINKLFGFRVKNW 551
DB 483 LVQLDSANGIIDVNSGKAKTQHILFRPLITPGTEHRRVQTQKYEYITKLNINRVDSW 542
QY 552 QVTDGEASSKLDFSKVLQV-----AETEGTDEIGLIVNAKAGNDDIFVGOQKMNIDG 604
DB 543 KITDGAASSTFDLTNVVQRIEILDNAGNVTKTKTIIIAKLGGEDDNNVFGSGTTEIDG 602
QY 605 GDGHDVRVYSGDGGFNIIVDGTSAEAGSYTVNRKVARGDIYHEVVVKRQETKVGKRTET 664
DB 603 GEGYDRVHYSR-GNYGALTIDATKETEGSSYTVNRFVETGKALHEVTSHTALVGNREK 661
QY 665 IOYRDYELRVKGYQSTDNLSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
DB 662 IEYR-HSNNQHHAGYTTKTLKAVEEIIGTSHNDIFKSGKFNDAFNGDGVDTIDGNDGN 720
QY 725 DRLFGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDNDDTLVDGTGNDKLA 784
DB 721 DRLFGKGGDDLLDGGNGDDFDGGKNDLHGGKGGDDIFVHRKGGDNDIITDSGNDKLS 780
QY 785 FADANISDIMIERTKEGIIIVKRNHDSGSIPIRWY----ITSNLQYQSNKTDHKIEOLI 840
DB 781 FSDSNLKDLPFEKVYKHLVI-TNSKKEKVTIONWFREADFAKEVPNYKATK-DEKIEII 838
QY 841 KGGSYITSDQIDKILQDKDGTVTISOELKKADENKSQKLSASDIASSLNKLVGSMAL 900
DB 839 GONGERITSQVDDLI--AKGNKGIQTODELSKVVDNYELLKHS-KNVTNSLDKLISVSVA 895
QY 901 FGTANSVSSNALQPIOTQGI 922
DB 896 FTSSNDSRNVLVAPTSMLDQSL 917

RESULT 5
US-07-908-253-2
; Sequence 2, Application US/07908253
; Patent No. 5534256
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: HARLAND, RICHARD J.
; TITLE OF INVENTION: HAEMOPHILUS SOMNUS OUTER MEMBRANE
; TITLE OF INVENTION: PROTEIN EXTRACT ENRICHED WITH IRON-REGULATED PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBERTA L. ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/908,253
; FILING DATE: 19920702
; CLASSIFICATION: 420
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9000-0026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
```


Db 721 DRLEGGKDDILDDGGDDFDGKGNLHGGKDDIFVHRKGDGNDIITDSDGNDKLS 780
Qy 785 FADANISDMIERKEGIIIVKRNHDSGSIINPRWY-----ITSNLQYQSNKTHKIEOLI 840
Db 781 FSDNLKDLTFEKKVHNLVI-TNSKKEKVTIQNFREADPAKEVPNYKATK-DEKIEEII 838
Qy 841 GKGSYITSDIOLKLODKDGTVITSOELKKLADENKSKLSASDIASSNLKLVGSMAL 900
Db 839 GONGERITSQVDDLI--AKNGKITQDELKSKVDVNYELLKHS-KVNTNSLDKLISSVSA 895
Qy 901 FGTANSVSSNALQPIQTQGI 922
Db 896 FTSSNDSRNVLVAPTSMLDQSL 917
RESULT 8
US-08-694-865-6
; Sequence 6, Application US/08694865
; Patent No. 5837268
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-694-865-6
Query Match 50.1%; Score 2326.5; DB 2; Length 926;
Best Local Similarity 50.0%; Pred. No. 1.5e-173;
Matches 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12;
Qy 15 LNSTKSGKLNLYLAI PKD--YDPOKGGTLDNFKAADDELGAELARLAERPNHETAKKSVDT 72
Db 7 LSFPFKGAKLILYIPQNYQVDTQGGQLQVLAEEGLIEVQREERNIATQTSIGT 66
Qy 73 VNQFLSLTQTGIAISATKLEKFLQKSTNKLAKGLDSVENIDRLKGRASNVLSLTSLFLG 132
Db 67 IQTAIGLTERGIVLSAPQIDKLQK---TKAGQALGSAESIVQNAKAKTVLSGIQSIIG 123
Qy 133 TALAGIELDSLKGGDAAPDALAKASIDLNEIIGNLSQSTQTTEATFSSQLAKLGSTISQ 192
Db 124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNLSLIENIANSVKTLDDEFGEIQSGKLN 182
Qy 193 AKGFSNIGNKQNL-NFSKTNLGLIEIITGLLSGISAGFALADKNASTGKKAAGFELSNO 251

Db 183 IKGLGTGLDKLKNIGLDRAGLGLDVISGLSGATAALVLAACKNASTAKKVGAGFELANQ 242
Qy 252 VIGNVTAKISSYVLAORVAAGLSTTGAVAAALITSSIMLAISPLAFNAADKPFHANALDE 311
Db 243 VVGNITKAVSSYTLAORVAAGLSTTGAVAAALITSSIMLAISPLAFNAADKPFHANALDE 302
Qy 312 FAKQPFKFGVGDHLLAEYORGVTIEASITTTISTALGASVAGSAAAVGSAVGTPIALL 371
Db 303 YAEFRKKGJGVDGNLAEYORGVTIDASTAINTALAAIAGGVSAAGSVIASPALL 362
Qy 372 VAGVTGLISGILEASKQAMFESVANRLOQKILEWEKQNGQNYFDKGYDSRYAAYLANNL 431
Db 363 VSGITGVISTILQYSKQAMFEHVANKIHNVKEWKNHGNKPNFYNGYDARYLANLQDNM 422
Qy 432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGRIKSKAGYADAPEDKKVEAGS 491
Db 423 KFLNLNKLQEAERVIAITQQRWDNNIGELAGITKLGRIKSKAGYADAPEDKKVEAGS 482
Qy 492 NITLDAKTGIIDISNSGKKTQALHFTSPLLTAGTESRRLTNGKYSYINKLKFRGVQW 551
Db 483 LVQDSANGIIDVNSGKAKTQHILPRLTTPGTEHREKRVQTKYEIITKLNINRVDWS 542
Qy 552 QVTDGEASSKLDPSKVIQV-----AETEGTDEICLIIVNAKAGNDIDFVGQGNIDG 604
Db 543 KITDGAASSTFDLTNVVQRIEIDNAGNVTKTKETIIAKLGEEDNVFVSGTTEIDG 602
Qy 605 GDGHRVYFSKGGPGNITVDGTSATEAGSYTVNRKVGDDIYHEVVKRQETKVGKRTET 664
Db 603 GEGYDRVHYSR-GNYGALITDATKETEQSGSYTVNRFVETGKALHEVTSTHTALVGNREK 661
Qy 665 IOYRDYELRKVGYGQSTDNKSVBEVIGSOENDFGSKFNDIFHSGEGDLDLDDGAGD 724
Db 662 IEYR-HSNQHHAGYITKTLKAVEEIIIGTSHNDIFGSKFNDAFNGGVDVTDIGNDGN 720
Qy 725 DRLEGGKGNDRLSGDEGDDLLDGGSGDDVNLGAGNDVYIFRKGDNNDTLYDGTGNDKLA 784
Db 721 DRLEGGKGNDRLSGDEGDDFDGKGNLHGGKDDIFVHRKGDGNDIITDSDGNDKLS 780
Qy 785 FADANISDMIERKEGIIIVKRNHDSGSIINPRWY-----ITSNLQYQSNKTHKIEOLI 840
Db 781 FSDNLKDLTFEKKVHNLVI-TNSKKEKVTIQNFREADPAKEVPNYKATK-DEKIEEII 838
Qy 841 GKGSYITSDIOLKLODKDGTVITSOELKKLADENKSKLSASDIASSNLKLVGSMAL 900
Db 839 GONGERITSQVDDLI--AKNGKITQDELKSKVDVNYELLKHS-KVNTNSLDKLISSVSA 895
Qy 901 FGTANSVSSNALQPIQTQGI 922
Db 896 FTSSNDSRNVLVAPTSMLDQSL 917
RESULT 9
US-08-694-865-6
; Sequence 6, Application US/08878748
; Patent No. 5969126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUI P.A.
; TITLE OF INVENTION: GRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS


```
Qy 73 VNQFLSLTGTGIAISATKLEFLQKSTNKLAKGLDSVENIDRKLGKASNVLSLSSFLG 132
Db 67 IQTAIGLTERGIVLSAPQIDKLLQK---YDPQKGGTLDNFTKAADELGIARLAERPNTETAKKSVDT 123
Qy 133 TALAGIELDSLIIKGGDAAPDALAKASIDLINEIIGLSQSTQTTIEAFSSQLAKGLGSTISQ 192
Db 124 SVLAGMDLDEAL-QNNSQHAKAGLELNTSLNIENIANSVKTLDERGEQISQFGSKLQN 182
Qy 193 AKGFSNIGKQLNL-NFSKTNLGLIITGLSGISAGFALADKNASTGKVAAGFELSNO 251
Db 183 IKGLGTLDGLKNIIGGLDVGSLGSLGATAALVLADKNASTAKKVGAGFELANQ 242
Qy 252 VIGNVTKAISVYLAQRAAGLSTTGAAVALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db 243 VVGNITKAVSSYILAQRAAGLSTGPAVALIASTVLSLAISPLAFAGIADKFNHAKSLES 302
Qy 312 FAKQFRKFGYDGDHLLAEYQYRGVGTIEASLTITSTALGAVSAGVAAVSGAVGTPIAL 371
Db 303 YAEFRKGLGYDGDNLLAEYQYRGVGTIDASVTAINTALAAIAGGVSAAGSVIASPIALL 362
Qy 372 VAGVTGLISGLEASKQAMPESVANRLOGKILEWEKQNGQNYFDKGYDSRYAYLANNL 431
Db 363 VSGITGVITSTLQYSQAMPEHVAHNKIHKIVWEKKNHGKNYFENGVDARYLANLQDNM 422
Qy 432 KFLSELNKELEAERVIAITQORWNNITGELAGITKLERIKSGKAYADAPEDGKKVEAGS 491
Db 423 KFLNLNKELEAERVIAITQOOWNNIGDLAGISRLGEKVLUSKAYVDAFEEGRHAKDK 482
Qy 492 NITLDAKTGIIDISNSGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLFGKRVKNW 551
Db 483 LVQLDSANGIIDVNSGKAKTQHILFRPLTLPCTEHRERVQTKGYEYITKLNINRVDSW 542
Qy 552 QVTGGEASSKLPFSKVRQV-----AETGTEIGLIVNAKAGNDDIFVQGGKQWIDG 604
Db 543 KITDGAASSTPDLTNVQVRIEGLDNAGNVTKTKETKIIAKLGEGLDNDVFGSGTTEIDG 602
Qy 605 GDGHRDVFYSKDGFGNITVDGTSATAGSVTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db 603 GEGYDRVHYSR-GNYGAUTIDATKETEQGSYTVNRFVETGKALHEVTSHTALVGNREEK 661
Qy 724 IQYRDYELRKVGYGYSQSTDNLSKVEEIVGSGFNDVFGSKFNDIFHSGEGLDLDGGAGD 724
Db 662 IEYR-HSNQHAGYTKTKLKAEEIIGTSHNDIFKSGKFNDAFNGGVDYTDGNDGN 720
Qy 725 DRLFGKGNDRSGDEGLDGGSDVLANGAGNDVYIPRKGDGNDTLYDGTGNDKLA 784
Db 721 DRLFGKGGDDI LDGGGDDFDGKGNDLLHGGKGGDDIFVHRKGGDNDIITDSDGNDKLS 780
Qy 785 FADANISDIMITERTKEGIIVKRNHSGSINIPRYV---ITSNLQYOSNKTDKHIEOLI 840
Db 781 FSDNLKDLTTEKVRKHNLI-VTNSKKEKVTIQNFWREADFAKEVFNKATK-DEKIEEII 838
Qy 841 KGDSYITSDQIDKILQDKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db 839 GONGERITSQVDDLI--AKNGKITQDELSKVVDNYELLKHS-KVNTNSLDKLISSVSA 895
Qy 901 FGTANSVSSNALQIPTQPTQGI 922
Db 896 FTSSNDSNRNLVAPTSMLDQSL 917
```

RESULT 11

US-09-124-491-6

; Sequence 6, Application US/09124491

; Patent No. 6022960

; GENERAL INFORMATION:

; APPLICANT: FOTTER, ANDREW A.

; APPLICANT: MANN, JOHN G.

; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: REED & ROBINS LLP

; STREET: 285 HAMILTON AVENUE, SUITE 200

```
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,491
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE: 09-AUG-1996
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-124-491-6
```

Query Match

50.1%; Score 2326.5; DB 3; Length 926;

Best Local Similarity 50.0%; Pred. No. 1.5e-173;

Matches 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12;

```
Qy 15 LNSTKSLKNLXLAIPKD--YDPQKGGTLDNFTKAADELGIARLAERPNTETAKKSVDT 72
Db 7 LSPFKTGAKKIILYIPQNYQYDTEQNGLODLVAAAEELGIEVQREERNIATAQTSIGT 66
Qy 73 VNQFLSLTGTGIAISATKLEFLQKSTNKLAKGLDSVENIDRKLGKASNVLSLSSFLG 132
Db 67 IQTAIGLTERGIVLSAPQIDKLLQK---YDPQKGGTLDNFTKAADELGIARLAERPNTETAKKSVDT 123
Qy 133 TALAGIELDSLIIKGGDAAPDALAKASIDLINEIIGLSQSTQTTIEAFSSQLAKGLGSTISQ 192
Db 124 SVLAGMDLDEAL-QNNSQHAKAGLELNTSLNIENIANSVKTLDERGEQISQFGSKLQN 182
Qy 193 AKGFSNIGKQLNL-NFSKTNLGLIITGLSGISAGFALADKNASTGKVAAGFELSNO 251
Db 183 IKGLGTLDGLKNIIGGLDVGSLGSLGATAALVLADKNASTAKKVGAGFELANQ 242
Qy 252 VIGNVTKAISVYLAQRAAGLSTTGAAVALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db 243 VVGNITKAVSSYILAQRAAGLSTGPAVALIASTVLSLAISPLAFAGIADKFNHAKSLES 302
Qy 312 FAKQFRKFGYDGDHLLAEYQYRGVGTIEASLTITSTALGAVSAGVAAVSGAVGTPIAL 371
Db 303 YAEFRKGLGYDGDNLLAEYQYRGVGTIDASVTAINTALAAIAGGVSAAGSVIASPIALL 362
Qy 372 VAGVTGLISGLEASKQAMPESVANRLOGKILEWEKQNGQNYFDKGYDSRYAYLANNL 431
Db 363 VSGITGVITSTLQYSQAMPEHVAHNKIHKIVWEKKNHGKNYFENGVDARYLANLQDNM 422
Qy 432 KFLSELNKELEAERVIAITQORWNNITGELAGITKLERIKSGKAYADAPEDGKKVEAGS 491
```

| | | | | | | |
|-----|-----------------------------|--------------------------|---------------|----------------|-------------|-----|
| 423 | KFLNLNKLQBARVIAITQOOWDNNI | GDLAGISRLGEKVL | SGKAYVDAFE | EGHKIAKDK | 482 | |
| 492 | NITLDAKTGIIDISNSNGKTKOALHFT | SPLLTAGTESERLITNGKYSINKL | FGVRQXW | 551 | | |
| 483 | LVLQDSANGIIDVNSGKAKTQHILFRT | PLLTPGTEHRRVQTKYEYITKLM | NRVDSW | 542 | | |
| 552 | QVTDGEASSKLDPSKVIQRV | -----AETGTD | IEGLIVNAKAGND | IPVGOKNNIDG | 604 | |
| 543 | KITDGAASSTFDLTNWQRIGIELD | NAGNVTKTKTIIA | KLGBGDDNV | FGSGTTEIDG | 602 | |
| 605 | GGCHDRVFSKDGFGFENITVDGTS | ATEAGSYTNVRKVARGD | IVHEVVKROETV | VGKRTET | 664 | |
| 603 | GEgyDRVHYSR-GNYGAULTIDAT | KETEQGSYTVNRFVETG | KALHEVTSHT | TALVGNREEK | 661 | |
| 665 | IQRDYELRKVGYQSTDNILKSVE | EVIGSGFNDVFKGSKFNDI | PHSGEGD | LLDGGAGD | 724 | |
| 662 | IEYR-HSNQHGAHYTTKTLKAVE | IIIGTSHNDI | FKGSKFNDA | PNFGDGVDTIDG | 720 | |
| 725 | DLRFGKGNDRLSGDGEDD | LLDGGSDDLVANGAGNDV | YIFRKDGDND | LYDGTGNDKLA | 784 | |
| 721 | DLRFGKGDDIILDGGNGDD | FDGCKGNDLLHGGKDDI | FVHRKGDGND | IIITDSDGNDKLS | 780 | |
| 785 | FADANISDI | MIERTKEGIVKRNDRHSGS | INIPRWY---- | ITSNLQNYQSNKTD | PHKIEQI | 840 |
| 781 | FSDNLKDLTTEKVKVHNLVI | -TNSKKEKVTTIQNN | FREADFAKEVPNY | KATK-DEKIEEII | 838 | |
| 841 | KDGGSYITSDOIKTLQDKDG | TWTSOELKKLADENK | SOKLSASDI | ASSINKLVGSMAL | 900 | |
| 839 | GQNGERITSKQVDDLLI | --AKNGKIKITQBELSK | VNDIYELLKHLS | --KXVINS | LDKLKJISVSA | 895 |
| 901 | FGTVANSVSSNALQPIQT | PQTQGI | 922 | | | |
| 896 | FTSSNDSRNVLVAPT | SMLDOSL | 917 | | | |

RESULT 12

US-09-383-912-6
; Sequence 6, Application US/09383912
; Patent No. 6521746
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, Thomas P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 amino acids

[illegible]

RESULT 15

RESOL 13
US-08-455-970A-10
; Sequence 10, Application US/08455970A

AGENCY NO.: 5706855
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HUW P.A.
TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
TITLE OF INVENTION: CHIMERAS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CALIFORNIA

| | | | | |
|----|--|-----|---|-----|
| Qy | | 605 | GDGHRVIFYSKDGFNGINTVDGTSGATEAGSYTVNRKVARGGDIYHEWVRQETKVGCRTET | 664 |
| Db | | 603 | GEGYDRVHYSR - GNYGALTIIDATKETEQSGYTVNRFVETGKALHEVTSTHTALVGNRBEK | 661 |
| Qy | | 665 | IQRDYELRKVGICYQSOTDNLKSVEEVITGSQFNDFPKSGKNNDI FHSGECDLLDGAGD | 724 |
| Db | | 662 | IERY- HSNQHAGYYTKTILKAVEEII GTSHNDIFPKSGKNDAFGGDDGVDTIDGNPQN | 720 |
| Qy | | 725 | DLRFGKGNDRLSGDEGBDDLDCGSGDDVLNCGAGNDVYIFRKGBCNDTLVDCGTGNDKLA | 784 |
| Db | | 721 | DLRFPGGKGDDILLDGGNGDDFIDGGKGNDDLHGCKGDDIFVIRKGCUNDNIITDSGNDKLS | 780 |
| Qy | | 785 | PADANISIMIERTEKEGIIVKRNDHSGSINIPRWY----ITSNLQONYOSNKTDHKIEQLI | 840 |
| Db | | 781 | FSDSNLKDLTPEVKYKHNLV-I-TNSKKKVTIQNWPFREADFAKEVENYKATK-DEKIEEII | 838 |
| Qy | | 841 | GKOGSYITSDQIDKILQDKDGTWITSQELKKLADENKSQKLASDISAASNKLUVGSNAL | 900 |
| Db | | 839 | GONGERITSKQVDDLI--AKNGKIQTELKSVVDNYELLKHS-KNVNTSLDKLITSSVSA | 895 |
| Qy | | 901 | PGTANSVSSNALQIPTQTQGI | 922 |
| Db | | 896 | FTSNSDRNVLVAPTSMLDQSL | 917 |

RESULT 16

```

US-08-455-970A-14
; Sequence 14, Application US/08455970A
; Patent No. 5708155
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
; TITLE OF INVENTION: CHIMERAS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,970A
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 951 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-970A-14

```



```

; MOLECULE TYPE: protein
; US-08-694-865-8

Query Match      50.1%; Score 2326.5; DB 2; Length 977;
Best Local Similarity 50.0%; Pred. No. 1.7e-173;
Matches 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12;

Qy 15 LNSTKSLGNLYLAIPKD--YDPQKGGTINDFTKAADELGIARLAEPNHTETAKSVDT 72
Db 7 LSPFKTGAKKILYIPONYQYDTEQNGLODLVKAABELGIEVQREERNNIATAQTSLGT 66

Qy 73 VNQFLSITQTGIAISATKLEKFLQKSTNKLAKGLDSVENIDRKLKASNVLSLTSFLG 132
Db 67 IQTAIGLTERGIVLSAIPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIOSILG 123

Qy 133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIGNLSQSTQTIAPSSOLAKLGSTISQ 192
Db 124 SVLAGMDLDEAL--QNNSNQHALLAKAGLELNTSLNIENIANSVKTLDFEGEQISQFGSKLQN 182

Qy 193 AKGFSNIGNKLQNL-NFSKTNLGLIEITGLLSGISAGFALADKNASTGKVAAGFELSNO 251
Db 183 IKGLGTIGDKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy 252 VIGNVTKAISSYVLAORVAAGLSTTGAAVALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db 243 VVGNITKAVSSYILAORVAAGLSTTGAAVALITSSIMLAISPLAFMNAADKFNHANALDE 302

Qy 312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASTITSTALGAVSAGVSAAVGSGVGTPIALL 371
Db 312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASTITSTALGAVSAGVSAAVGSGVGTPIALL 371

US-08-878-748-8
; Sequence 8, Application US/08878748
; Patent No. 5969126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GORH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,748
; FILING DATE: 19-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-878-748-8

Query Match      50.1%; Score 2326.5; DB 2; Length 977;
Best Local Similarity 50.0%; Pred. No. 1.7e-173;
Matches 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12;

Qy 15 LNSTKSLGNLYLAIPKD--YDPQKGGTINDFTKAADELGIARLAEPNHTETAKSVDT 72
Db 7 LSPFKTGAKKILYIPONYQYDTEQNGLODLVKAABELGIEVQREERNNIATAQTSLGT 66

Qy 73 VNQFLSITQTGIAISATKLEKFLQKSTNKLAKGLDSVENIDRKLKASNVLSLTSFLG 132
Db 67 IQTAIGLTERGIVLSAIPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIOSILG 123

Qy 133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIGNLSQSTQTIAPSSOLAKLGSTISQ 192
Db 124 SVLAGMDLDEAL--QNNSNQHALLAKAGLELNTSLNIENIANSVKTLDFEGEQISQFGSKLQN 182

Qy 193 AKGFSNIGNKLQNL-NFSKTNLGLIEITGLLSGISAGFALADKNASTGKVAAGFELSNO 251
Db 183 IKGLGTIGDKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy 252 VIGNVTKAISSYVLAORVAAGLSTTGAAVALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db 243 VVGNITKAVSSYILAORVAAGLSTTGAAVALITSSIMLAISPLAFMNAADKFNHANALDE 302

Qy 312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASTITSTALGAVSAGVSAAVGSGVGTPIALL 371
Db 312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASTITSTALGAVSAGVSAAVGSGVGTPIALL 371

US-08-878-748-8
; Sequence 8, Application US/08878748
; Patent No. 5969126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GORH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,748
; FILING DATE: 19-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-878-748-8

```

```
Db 303 YAEFFKGLYDGNLLAEYQRTGTIDASTAINTALAAAGVSAAGSVAIPIALL 362
Qy 372 VAGVTGLISGLEASQAMFESVANRLQKILEWQKQNGQYFDKGYDSRYAAYLANNL 431
Db 363 VSGITGVISTILQYSQAMFHVANKIHKIVEKNNHGNYPENGVDARYLANLQDNM 422
Qy 432 KFLSELNKELEAERVAITQORWNNIGELAGITKGERIKSGKAYADAFEDGKKVAGS 491
Db 423 KFLNLNKLQAEARVAITQQQWNNIGDLGAGISRLGKVLVSQKAYDAFEEGKHAKD 482
Qy 492 NITLDAKTIIDISNNGKKTQALHFTSPLLTAGTESRLTNGKYSYINKLKEFRVKNW 551
Db 483 LVQLDSANGIIDVNSNGKAKTQHILFRTPLTPGTEHREVRVQTKYEVITKLNINRVD 542
Qy 552 QVTDEASSKLDKFSKVIQV-----AETEGTDEIGLIVNAKAGNDDIFVGGQNMIDG 604
Db 543 KITDGAASSTFDLTNNVQRIEIDNAGNVTKTETKIIAKLGEEDDNNVFGSGTTEIDG 602
Qy 605 GDGHDVRYSKDGGFNGNITVGTSTAGTSATAGSVTVNRKVARGDIYHEVVKRQETKVKRTET 664
Db 603 GEGYDRVHYSR-GNYGALTIDATKETEQQSVTVNRVETGKALHEVTSHTALVGNREEK 661
Qy 665 IQYRDYELRKVGYQSTDNLSVEEVIGSQFNDVFGSKFNDIFHSGEGDLDLGGAGD 724
Db 662 IEYR-HSNNHAGYYTKTLKAVEEIICTSHNDIFKSGKFNDAFNGDGVDTIDGNDGN 720
Qy 725 DRLFGKGNDRLSGDEGDDLDGGSGDVLNGAGNDVYIFRKGNDNTLYDGTGNKLA 784
Db 721 DRLFGKGGDLDLGGGDDFDGKGNLHGGKGGDIFVHRKGGDNDIITDSGDNKLS 780
Qy 785 FADANISDIEMTERTKEGIIKVRNDHSGSINIPRY-----ITSNLQYOSKTDHKEOLI 840
Db 781 FDSNKLKDLTEKVKHNLVI-TNSKKEKVTIQNMFREADFAKEVPNYKATK-DEKIBEI 838
Qy 841 GKDGSIYTSDDIKLQKDDGTVITTSQELKADENKSKLSASDIASSLNKLGVSMAL 900
Db 839 QONGERITSKQVDDLI--AKNGKITQDELSKVVDNYELLKHS-KVNTSLDKLISVSA 895
Qy 901 FGTANSVSSNALQPTQPTQGI 922
Db 896 FTSSNDRNLVAPTSMLDQSL 917
```

RESULT 20

US-09-124-491-8

Sequence 8, Application US/09124491

Patent No. 6022960

GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.

APPLICANT: MANN, JOHN G.

TITLE OF INVENTION: GRH-LEUKOTOXIN CHIMERAS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESS: REED & ROBINS LLP

STREET: 285 HAMILTON AVENUE, SUITE 200

CITY: PALO ALTO

STATE: CA

COUNTRY: USA

ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09124,491

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/694,865

FILING DATE: 09-AUG-1996

APPLICATION NUMBER: US 08/387,156

```
Qy 15 LNSTKSGKMLYLAIKPD--YDQKGGTINDPFIKAADELGIARLAEPNHTETAKKSVDT 72
Db 7 LSPFKTGAKKILYIPQNYQYDTEQNGLODLVKAABEELGIEVQREERNIATQTSLTG 66
Qy 73 VNQFLSTQGTGIAISATKLEKFLQKSTNKLAKGLDSVENIDRKLKASNVLSLSSFLG 132
Db 67 IQTAIGTERGIVLSAPQIDKLQK---TKAGQALGSAESIVQNVANKAKTVLSQISILG 123
Qy 133 TALAGIELSLIKGDAAPDALAKASIDLINEIGNLSQSTQTIEAFSSQAKLGGSTISQ 192
Db 124 SVLAGMDLDEAL-QNNSNQHALAKAGLELNTSLNIENIANSVKTLDEFGEQISQFGSKLQN 182
Qy 193 AKGESNIGNKQNL-NFSKTNLGLIITGLISGAFALADKNASTGKVKVAGFELSNO 251
Db 183 IKGLTGLDKLKNIGGLDKAGLDGLVIGLSGATAALVLADKNASTAKKVGAGFELANQ 242
Qy 252 VIGNVTKAISSYVLAQRAAGLSTTGAAALITSSIMLAISPLAFMAADKFHNAALDE 311
Db 243 VGNITKAVSYILAQRAAGLSTGPAVALIASTVSLAISPLAFAGIADKFHNAKLES 302
Qy 312 FAKQFRKFGYDGHLLAEYQRTGTIDASTAINTALAAAGVSAAGSVAIPIALL 371
Db 303 YAEFFKGLYDGNLLAEYQRTGTIDASTAINTALAAAGVSAAGSVAIPIALL 362
Qy 372 VAGVTGLISGLEASQAMFESVANRLQKILEWQKQNGQYFDKGYDSRYAAYLANNL 431
Db 363 VSGITGVISTILQYSQAMFHVANKIHKIVEKNNHGNYPENGVDARYLANLQDNM 422
Qy 432 KFLSELNKELEAERVAITQORWNNIGELAGITKGERIKSGKAYADAFEDGKKVAGS 491
Db 423 KFLNLNKLQAEARVAITQQQWNNIGDLGAGISRLGKVLVSQKAYDAFEEGKHAKD 482
Qy 492 NITLDAKTIIDISNNGKKTQALHFTSPLLTAGTESRLTNGKYSYINKLKEFRVKNW 551
Db 483 LVQLDSANGIIDVNSNGKAKTQHILFRTPLTPGTEHREVRVQTKYEVITKLNINRVD 542
Qy 552 QVTDEASSKLDKFSKVIQV-----AETEGTDEIGLIVNAKAGNDDIFVGGQNMIDG 604
Db 543 KITDGAASSTFDLTNNVQRIEIDNAGNVTKTETKIIAKLGEEDDNNVFGSGTTEIDG 602
Qy 605 GDGHDVRYSKDGGFNGNITVGTSTAGTSATAGSVTVNRKVARGDIYHEVVKRQETKVKRTET 664
Db 603 GEGYDRVHYSR-GNYGALTIDATKETEQQSVTVNRVETGKALHEVTSHTALVGNREEK 661
Qy 665 IQYRDYELRKVGYQSTDNLSVEEVIGSQFNDVFGSKFNDIFHSGEGDLDLGGAGD 724
Db 662 IEYR-HSNNHAGYYTKTLKAVEEIICTSHNDIFKSGKFNDAFNGDGVDTIDGNDGN 720
```

Query Match 50.1%; Score 2326.5; DB 3; Length 977;
Best Local Similarity 50.0%; Pred. No. 1.7e-173;
Matches 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12;

```

Qy 725 DRLFGGKGNDRSLGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGNDGNDLYDGTGNDKLA 784
Db 721 DRLFGGKGNDRSLGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGNDGNDLYDGTGNDKLA 780
Qy 785 PADANISDIMEIKTEGIIIVKRNHSGSINIPRWY----ITSNLQNYOSNKTIDHKIBOLI 840
Db 781 FSDSNLKDLPFEKVKNLVI--TNSKKEKVITQNFREADFAKEVPNYKATK-DEKIEI 838
Qy 841 GKDSYITSDQIDKILQDKKDGTVITTSQELKKLADENKSKLSASDIASSLNKLVGSMAL 900
Db 839 GONGERITSKQVDDLI--AKNGKITQDELKSKVDNYELLKHS--KNVTNSLDKLISVSA 895
Qy 901 FGANSVSSNALQPIPTQTOGI 922
Db 896 FTSSNDSRNVLVAPTSMLDQSL 917

RESULT 21
US-09-383-912-8
; Sequence 8, Application US/09383912
; Patent No. 6521746
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GARH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-383-912-8

Query Match 50.1%; Score 2326.5; DB 4; Length 977;
Best Local Similarity 50.0%; Pred. No. 1.7e-173;
Matches 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12;

Qy 15 LNSTKSGULNLYLAIPKD--YDPQKGGTLDNFKADELGIARLAEEPNTHTAKSVYDT 72
Db 7 LSPFKTGAKIILYIPQNYQYDTEGGNGLDLVKAABELGIEVQREERNNTAQTSLGT 66
Qy 73 VNQFSLTQTGTGATGATKLEKFLQKHSNKLAKGLDSVNTDRKLGKASVNTSLSSFLG 132
Db 67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGALGSAESIVQNAKAKTYLSIGTSLG 123

```

RESULT 22

US-07-777-715-9

; Sequence 9, Application US/07777715

; Patent No. 5273889

; GENERAL INFORMATION:

; APPLICANT: Potter, Andrew

; APPLICANT: Campos, Manuel

; APPLICANT: Hughes, Huw P.A.

; TITLE OF INVENTION: CYTOKINE-LEUKOTOXIN GENE FUSIONS AND

; TITLE OF INVENTION: USES THEREOF

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Morrison & Foerster

; STREET: 545 Middlefield Road, Suite 200

; CITY: Menlo Park

; STATE: California


```
Qy 492 NITDAKTGIIIDNSNGKTKQALHFTSPILTTACTESRLTNGKYSYINKLKFGKVKW 551
Db 483 LVQLDSANGIIDVNSNGKAKTOHILFRTPLTPGTEHRRVQTGKYEYITKLNINRVDSW 542
Qy 552 QVTDGEASSKLDVSKVIQV-----AETEGTDEIGLIVNAKAGNDDIFVGOQKMNIDG 604
Db 543 KITDGAASSTFDLNNVQRIEGLDNAGNVTKTETKIIAKLGBGDDNVFVSGTTEIDG 602
Qy 605 GDGHRVPYKDGFGNITVDGTGATEAGSVTVNRKVARGDIIYHEVVKRQETKVGKRTET 664
Db 603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSHTALVGNREEK 661
Qy 665 IOYRDYELARKVGYQVOSTDNLSKVEEIVGSQFNDVFKSGKENDIFHSGEGDLDLDGAGD 724
Db 662 IEYR-HSNQHAGYITKDTLKAVEEIIIGTSHNDIFPKSGKENDAFNGDGGVDTIDGNDGN 720
Qy 725 DRLFGKGNDRLSGDEGDDLDGSGDDVNLGAGAGNDVYIIPRKGDDGNDTLVDGTGNDKLA 784
Db 721 DRLFGKGDDLDGSGDDDFDGGKGNDDLHGKGGDDIFVHRKGGDNDIITDSGNDKLS 780
Qy 785 FADANISIMIERKKEGIIIVKRNHDSGSINIPRWY----ITSNIONYOSNKTDKHIBOLI 840
Db 781 FSDSNLKDLPKVKHNLVI-TNSKKEKVTIQQNWFREADFAKEVPNYKATK-DEKIEBII 838
Qy 841 GKDGSIITSDQIDKLQDKDGTVITSOELKLADENKSQKLSASDIASSLNKLVGSNAL 900
Db 839 GQNGERITTSKVDDLI--AKNGKIKITQDELSKVDNVELLKHS-KNVTNSLKLSSVSA 895
Qy 901 FGTANSVSSNALQIPITQPTQGI 922
Db 896 FTSSNDSRNLVAPTSMLDQSL 917
```

RESULT 25

```
US-08-215-805A-80
; Sequence 80, Application US/08215805A
; Patent No. 5559008
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; TITLE OF INVENTION: LEUKOTOXIN GENE FROM PASTEURBELLA
; TITLE OF INVENTION: SUITS
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/215.805A
; FILING DATE: 22-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/61 (D-1329A)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
; ORIGINAL SOURCE:
; ORGANISM: Pasteurella suis
; STRAIN: 5943
; IMMEDIATE SOURCE:
; LIBRARY: P. suis DNA in Bacteriophage lambda-dash
; CLONE: (lambda)yfc33-37
; US-08-215-805A-80
```

Query Match 49.4%; Score 2297.5; DB 1; Length 934;

Best Local Similarity 49.7%; Pred. No. 2.9e-171; Indels 49; Gaps 17;
Matches 469; Conservative 172; Mismatches 254;

```
Qy 1 MSNINV-IKSNIOAGLNST-----KSLGNLNYLAIPKD--YDPQKGGTFLNDFIK 46
Db 4 LANISTNLKNSLOGLHKTQSLNQAQSLKAGAKKILYIPKDYEDSGRGNGLQDLVK 63
Qy 47 AABELGTARLAEBPNHTETAKKSVDTVNQPLSTQTGTGIAISATKLEKFLQKHSNKLAKG 106
Db 64 AEDDLGIEVQREERNGIATQAQNSLSTITNTLGFSEGVLSAPQLDKLQKY---KISKA 120
Qy 107 LDSVENIDRKLKGASNVLSLSSFLGTALAGIELDSLIIKKGDAAPDALAKASIDLINELI 166
Db 121 PGSEENVAKNLGNAQTLLSGIQSILGSVMAGMDLDELKKNKSGELD-LAKAGLELTNLSLI 179
Qy 167 GNLSQSTQTTEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-FSKTLMGLEIITGLLSGI 225
Db 180 ENTANSVQITLDTFSEQLSGTGLKQNVKGLTGLDKLKNFSGSKAGLGLLEVISGLLSGA 239
Qy 226 SAGFALADKNASTGKVAAGFELSNOVIGNVTKAISYVLAQRAAGLSTTGVAALITS 285
Db 240 TAALVLADKNASTDRKVGAGFELANQVGNITKAVSSYIIAQRVAAGLSNTPVPSALIAS 299
Qy 286 SIMLAISPLAFMAADKFNHANALDEFAKOFKRGVGDGHLLEAYORGVGTIEASLTIS 345
Db 300 TVALATSPFAPGATDKFNNAKALSYAERFKLGYEGDSLLABYQGTGTIDASVTAVN 359
Qy 346 TALGAVSAGVSAAVGSAVGTPIALLVAGVTGLISGILEASKQAMPFESVANRLQKILEW 405
Db 360 TALAAISGGVSAAGSLVGAPIALLVSGITGIITILQYSKQAMFEHVANKIHKIVDM 419
Qy 406 EKQNGGNYFDKGYDSRYAAYLANNLKFLSELKLEBAERVIAITQORWNNIGELAGIT 465
Db 420 EKHNGKNYFENGYSRYLADLQNMRLQNLKLEOAERVIRITQOQWNNIGNLAGIS 479
Qy 466 KLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTIIDISNKGKTKQALHFTSPILTAG 525
Db 480 RLGEKVMGSKAYADAFEEGKLIRADTFVQDLSATGVTNKSNNVKTQHILFRTPLTPG 539
Qy 526 TESRERLTNGKYSYINKLKFGKVKNMQVTDGEASSKLDVSKVIQV-----ETEGTD 578
Db 540 VENRERIQTGKYEYITKLNINRVDSWKITDGTATNSTFDLTNNVQRIEGLDHADNVTKTK 599
Qy 579 EIGLIVNAKAGNDIIFVGOQKMNIDGSDGHRVPYKDGFGNITVDGTGATEAGSVTVN 638
Db 600 ETKIIANLGDNDVDFIGSGTTEVDGNGGLDRVHYSR-GDYGALTIDATNESVQGSYTVK 658
Qy 639 RYVARGDIIYHEVVKRQETKVGKRTETIOYRDYELARKVGYQVOSTDNLSKVEEIVGSQFND 698
Db 659 RFVETGKALHEVTATQSVLVGSREEKIEYR-HSNNTQAHAGYITDTLKSVEEIIIGTGRND 717
Qy 699 VFKGSKFNDFIHFSGEGDLDLDGAGDDRLFGKGNDRLSGDEGDDLDLDGSGDDVNLGAGG 758
Db 718 IFKSGKFDFAFHGGDGVNDIDGNAGNDRLFGKGFDDIIDGGDGDFFIDGGGDDILHGK 777
Qy 759 GNDVYIIPRKGDDNTLYDGTGNDKLAIFADANISIMIERKKEGIIIVKRNHDSGSINIPRW 818
Db 778 GNDILCTVKG-GNDSISDSGNDRLSFADSNLKDLPKVKHNLMI-TNVKKEKVTIQNW 835
Qy 819 Y-----ITSNLYOSNKTDKHIEOLICKDSYITSDQIDKLQDKDGTVITSOELKLA 874
Db 836 FREADYAKTVHNYQAT-ADEKIEIIGRQGERITTSKQIDELIEKKG--KIDQSELERIA 892
Qy 875 DENKSQKLSASDIAS-SLNKLVGSMALFGTANSVSSNALQIPITQ 917
```



```
Best Local Similarity 47.7%; Pred. No. 1.8e-166;
Matches 453; Conservative 188; Mismatches 255; Indels 53; Gaps 15;

Qy 1 MSNINV--IKSNIQAGL-----NSTKSGKLNLYLAIPKDYDPQ 36
Db 1 MSKITLSSLKSSLOQLANGKNKLNQAGTTLKNGLTQTGHSLQNGAKKLILVIFQGYDSG 60

Qy 37 KGGTLDNFIKADELIGIARLAEPNHTETAKKSVDTVNOFSLTQTGTGIAISATKLEKPLQ 96
Db 61 QGNGVQDLVKAANDLIGIEVWEERSNLDIAKTSFDTTQKILGFTDRGIVLFPAPOLDNLK 120

Qy 97 KHSNTKLAKGDSVENIDRKLGKASNVLSLSPFLGTALAGIELDSLKKGDAAPDALAK 156
Db 121 KNP--KIGNTLGSASSISQNTGKANTVLGGQISLGSVLGVNLLQNKPQNOLEAK 178

Qy 157 ASIDLINIEIIGNLSTQTIETAFSSQLAKLSTISQAQFSGNIGNKLQNL-NFSKTNLGL 215
Db 179 AGLLELTNVLGNIASSVQTVDAFAEQISKLSHLQNVKGLGSLNKLQNLPLDKASLGL 238

Qy 216 EIITGLLSGISAGFALADKNASTGKKVAAAGFELSNOVIGNTVTKAISVVLAAQRAAGLST 275
Db 239 DIISGLSGASAGLILADKEASTEKAAGAAGVEFANQIIGNTVKAVSSVILAAQRAAGLSS 298

Qy 276 TCVAVALITSSIMLAISPLAFMNRADKENHANALDEFAPKRFKGYDGDHLLAEYQGVG 335
Db 299 TGPVAALASTVALAVSPLFSLFVNADKPKQADLIKSYERFKQUGYDGDRLADPHRETG 358

Qy 336 TIEASLTITISTALGAVSAGVAAAAGSVAGTPIALLVAGVTGLISGILEASKQAMFESVA 395
Db 359 TIDASVTITNTALAAISGVGAASAGSLVGPVALLVAGVTGLITILEYSKQAMFEHVA 418

Qy 396 NRLOQKILWEKQNGQNYFDKGYDSRYAAVYANNLKLSELNKELEAEVIAITQQRWD 455
Db 419 NKVDRIIVEWEKCH-KNKYFEQGYDSRHLADLQNMKFLINLKELOAEVVAITQQRWD 477

Qy 456 NNIGLAGITKGERIKSGKAYADAFEDGKVEAGSNITTLAKTGIIDISNSNGKKTQAL 515
Db 478 NOIGLAAISRDTKISSGKAYDAFAFEGQHQSYDSVQLDNKNGINISNTN-RKTQSV 536

Qy 516 HFTSPLLTAGTESRRLTNGKYSYINKLKFGRKVMQVTDGEAASSKLDKFSKVIQKVA--- 572
Db 537 LFRTELLTPGEENRERIQEGKNSYITKLHIQRVDSMTVTDGASSVDFTNVVQRIAVKF 596

Qy 573 -----ETEGDEIGLIVNAKAGNDIDFVGGKKNIDGGDHRVFKSKDGGFENITVD 625
Db 597 DDAGNIIIESKDTK---IIANLGAGNDNVFVGSSTTVIDGGDGHDRVHSR-GEYCALVID 652

Qy 626 GTSATEAGSYTVNRKVARGDIVHEVVKQETKVGKRTETIOYRDVELKRVGVGQSTDL 685
Db 653 ATAETEKSGYSVKRYVGDSSKALHETIATHQTNVGNREEKIEYR-REDDRFHTGTVTVTDSL 711

Qy 686 KSVEEVIGSQNDVDFKSGFNDIFHSGEGLDLDGGAGDDRLFGKGNDRLSGDEGDDLL 745
Db 712 KSVEEIIIGSQFNDIFKGSQFDDVFHGGNGVDTIDGDCDDHLFGAGDDVIDGGNGNFL 771

Qy 746 DGGSGDDVLNGAGNDVYIFRKGDNGLYDGTGNDKLAPADANISDIMIERTKEGIIVK 805
Db 772 VGGTGNDIISGGKNDIYVHKTGDNDSITDSGGQDKLAFSDVNLKOTFFKKVDSLSLEI- 830

Qy 806 RNDHSGSINIPRWY----ITSNLQYQSNKTDHKTEQLIGKDGYSYITSDQDKTLODKKD 861
Db 831 INQKEKVRIGNWFLDEDDLASTVANYKAT-NDRKLEIEIIGKGERITSEQVDKLI--KEG 887

Qy 862 GTVITSQBLKXLADENKSKLASDIASSLNKLVGSMALFGTANSVSN 910
Db 888 NNQISAEALSQVNDYNTSK-DRONVSNLAKLISSVGSFTSSSDFRN 935
```

RESULT 28

```
US-09-062-126-10
; Sequence 10, Application US/09062126
; Patent No. 6500435
; GENERAL INFORMATION:
```

```
; APPLICANT: Kamp, Elbarte Margriet
; APPLICANT: Smits, Marinus Adrianus
; TITLE OF INVENTION: Recombinant Vaccine For Prevention
; TITLE OF INVENTION: And/Or Treatment Of Pleuropneumonia Infections
; FILE REFERENCE: 470-980537
; CURRENT APPLICATION NUMBER: US/09/062,126
; CURRENT FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 08/488,706
; PRIOR FILING DATE: 1995-06-09
; PRIOR APPLICATION NUMBER: 08/138,609
; PRIOR FILING DATE: 1993-10-15
; PRIOR APPLICATION NUMBER: 07/722,971
; PRIOR FILING DATE: 1991-06-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1053
; TYPE: PRT
; ORGANISM: Actinobacillus pleuropneumonia
; US-09-062-126-10
```

```
Query Match 41.4%; Score 1926; DB 4; Length 1053;
Best Local Similarity 42.1%; Pred. No. 4.7e-142;
Matches 420; Conservative 168; Mismatches 290; Indels 120; Gaps 19;
```

```
Qy 8 KSNIQAGLNSTKSGLK-----NLYLAIPKDYDPQKGGTLDNFI 45
Db 19 KROAKKGYDVTGLOGLYGVYQAKLQALGAGKAVQKYGKLVIPKGYDGSVGNGFDDL 78

Qy 46 KADELIGIARLAEPNHTETAKKSVDTVNOFSLTQTGTGIAISATKLEKFLQKHSNKLAK 105
Db 79 KAABELGIVQYVNRNELEVANHKTGTADQFLGTERGLTLFAPQLQDFLQKHSKSNV 138

Qy 106 GLDSVENIDRKLGKASNVLSLSPFLGTALAGIELDSLKKGDAAPDALAKASIDLNEI 165
Db 139 GSSTGDAVS-KLAKSQTIISGISOVLGTLVAGINLEAIISSGSELE-LAEAGVSLASEL 196

Qy 166 IGNUSTQTIETAFSSQLAKLSTISQAQFSGNIGNKLQNL---FSKTNLGLIEITGLL 222
Db 197 LSNIAKGTITTDATFTTQIQNFGLVENAKLGGVGRQLQNISSALSKTGLGLDIISLL 256

Qy 223 SGISAGFALADKNASTGKKVAAAGFELSNOVIGNTVTKAISVVLAAQRAAGLSTGVAAL 282
Db 257 SGVTASFALANKVASTSTKVAAGFELSNOVIGITKAVSSVILAAQRAAGLSTGPAAL 316

Qy 283 ITSSIMLAISPLAFMNAADKENHANALDEFAPKRFKGYDGDHLLAEYQGVGVTIEASLT 342
Db 317 IASSISLAISPLAFLRVADNFRNSKEIGFAERFKLGYDGDKLLSEFYHEAGTIDASIT 376

Qy 343 TISTALGAVSAGVAAAAGSVAGTPIALLVAGVTGLISGILEASKQAMFESVANRLOKI 402
Db 377 TISTALSAIAAGTAAASAGALVGAPITLLVTGITGLISGILEFSKQPLDHVASKIGNKI 436

Qy 403 LEWKONGQNYDPKGYDSRYAAVYANNLKLSELNKELEAEVIAITQQRWDNNIGELA 462
Db 437 DEWEKKY-GKNYFENGDAHKAFLSDSFSLLSFNKOYETERAVLITQQRWDYIGELA 495

Qy 463 GITKLGBRIKSGAYADAFEDGKVKVEAG----SNITLDAKTIIDISNSNGKKTQALHFT 518
Db 496 GITKGDKLSSGKAYVDYFQEGKLEKPPDDFSKVPFDPKGEIDISNS--QTSTLLKVF 553

Qy 519 SPLITAGTESRRLTNGKYSYINKLKFGRKVMQVTDG--EASSKLDKFSKVIQK- 570
Db 554 TPLLTPTGESRERTQTGNYYEITKLVVYKDKW-VVNGVXDKGAVYDTNLIQAHHTSS 612

Qy 571 VAETEGTDEIGLIVNAKAGNDIDFVGGKKNIDGGDGHDRVFKSKDGGFNGITVDGTSAT 630
Db 613 VARGEYREVRLVSHLGNLGNDKVFLVAGSAEIHAGEGHDVYYDKT-DTGLLVIDGTKAT 671

Qy 631 EAGSYTVNRKVARG-DIYHEVVKRQETKVGKRTETIOYRDELKRVG-YGVQSTDNLKSV 688
Db 672 EQRYSVTBELSGATKILREVNIKQSAVGKREETLEYRDELQSGNSNLKAHDELHSV 731
```



```

; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 6580-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1022 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Actinobacillus pleuropneumoniae
; US-08-772-270A-2

Query Match. 37.9%; Score 1763.5; DB 3; Length 1022;
Best Local Similarity 40.0%; Pred. No. 2.4e-129;
Matches 400; Conservative 181; Mismatches 309; Indels 109; Gaps 26;

Qy 8 KSNIOAGLNTSKSL-----KNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
Db 21 KSAKSGAGALKNGLGQVKQAGQKLLIYPKDYQASTGSLNDLVKAAEALGIEVHRSEK 80
Qy 61 NHTETAKSVDTVNOFLSLTQTGIAISATKLEKFLQKHSNKLAKGL-DSVENIDRKLK 119
Db 81 NGTALAKELFGTTEKLLGFSERGIALFAPQDFKLKLNK--QKLSKSLGSSSEALGORL 138
Qy 120 ASNVLSLTSFLGTALAGIELDSLKK-----GDAAPDALAKASIDLINEIIGNLSOST 175
Db 139 TQTALSAQLGFLGTATAGMDLSLRRRNGEDVSGSELAKAGVDLAAQLVDMNIASATG 198
Qy 176 IEAFSSQLAKLGTISQAKGFSNIGNKLQNL-NFSKTNLGLIITIGLISGISAGFALADK 234
Db 199 VDAPAEQGLKGNALNTR-LSGLASKLNNLPDLSLAGPGFVDAVSGILSVVSASFILSNK 257
Qy 235 NASTGKKVAGFELSNOVGNVTIAISSYLAQVAAGLSTTGVAALITSSIMLAISPL 294
Db 258 DADAGTAAAGIEISTKILGNIGKAVSQYIIAQRVAAGLSTTAATGGLIGSVVVALAISPL 317
Qy 295 AFNNAADKFNHANALDEPAKQFKFGYDGHLLAEYQGVGTIEASLTITISTALGAVSAG 354
Db 318 SFLNVADKFERAKOLEQYSEKFKFGYDGSLLASFYRETGAIEAALITINSVLSAASAG 377
Qy 355 VSAAVGSAGVGTPIALLVAGTGLISGILEASKQAMFESVANRLQKILEWEKQNGQNY 414
Db 378 VGAATGSLVGPAAALVSAITGIISGLDASKQAI FERVATKLANKIDWEKXH-GKNY 436
Qy 415 PKGYDSRYAAYLANNLKFTSELNKELEAEKVIAITQORWNNIGELAGITKGERIKSG 474
Db 437 FENGYDARHSAFLDTEFELLSQYNKEYSVERVAITQORWNVNIGELAGITRKGSDTKSG 496
Qy 475 KAYADAEDGKGVKAG---SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRE 530
Db 497 KAYVDFEEGLKKEPRDPKVPDPLEGKIDUSSIN--KTTLLKVPVPTFAGEIRE 554
Qy 531 RLTKGKTSYINKLKFRGVQWQVTDGEASSKL--DFSKVIQVARTEGTDEIGLIVNAKAG 589
Db 555 RKQTKGYEYMTFLVKGEKQWVTGVQSHNAIYDTNLIQLAIDKKG-EKRQVITESHLG 613
Qy 590 --NDDIFVGQGMNDGGDGHDRVFSK-DGGFNGNITVDGTSATEAGSYTNRVK-ARGD 645
Db 614 EKNDRIVLSSGSSIVYAGNGHDVAAYDKTDGTY--LTFDQSOAKAGEIYVTKELKADVK 671
Qy 646 IYHEWVKQETKVKGRRTETIORYDELK--KVGYGYQSTONLKSVEEVIGSQFNDVFKGS 703
Db 672 VLKEVVKQTQDISVGRKSRSEKLEYRDELSPPELFGNIGIRAKDELHLSVEEILGSRNKKPF 731
Qy 704 KFNDFIHSGEED-----DLLOGGAGDRLFLGCKGNDRILSGDEGDDLL 745
Db 732 RFTDIFHAKGDDDEIYGNDDHDLIYGGDNDVTHGGDNDHLVCGNGNDRILGKGNFL 791

```

```

Qy 746 DGSGGDD-----VLNGGAGNDVY-----IFRKGDGNDTLYDGTGND----- 781
Db 792 NGGDGDELOVFEQGVNVLGGAGNDILYSGDGTNLFDGCVGNDKLYGLGKDIYRSKE 851
Qy 782 -----KLAADANI-----SDMIERTKEGIIVKRNDHSG-SINIPR 817
Db 852 YGRHIIIEKGGDDTLLSDLSFKDVGFRIGDGLLVNKRIGTGLYYHEDYNGNALTIKD 911
Qy 818 WYITSLQNYQSNKTDHKEIQLEIGKQGSYITSDQDKLQDKDGTGTVITSOELKLADEN 877
Db 912 WF--KEGKEGQNN---KIEKIVDKGAYVLQYLTETAPGRGINFYNGLEKFLYYGEG 965
Qy 878 KSKQLSASDIASSLNKLGVSMALF-CTANSVSSNALQPI 915
Db 966 YN---ALPQIRKDIQEIISSTGFTGDHKGKSVVSGGPL 1001

RESULT 32
US-09-062-126-3
; Sequence 3, Application US/09062126
; Patent No. 6500435
; GENERAL INFORMATION:
; APPLICANT: Kamp, Elbarte Margriet
; APPLICANT: Smits, Marinus Adrianus
; TITLE OF INVENTION: Recombinant Vaccine For Prevention
; TITLE OF INVENTION: And/Or Treatment Of Pleuropneumonia Infections
; FILE REFERENCE: 470-980537
; CURRENT APPLICATION NUMBER: US/09/062,126
; CURRENT FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 08/488,706
; PRIOR FILING DATE: 1995-06-09
; PRIOR APPLICATION NUMBER: 08/138,609
; PRIOR FILING DATE: 1993-10-15
; PRIOR APPLICATION NUMBER: 07/722,971
; PRIOR FILING DATE: 1991-06-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Actinobacillus pleuropneumonia
; US-09-062-126-3

Query Match 37.7%; Score 1751.5; DB 4; Length 1013;
Best Local Similarity 40.1%; Pred. No. 2.1e-128;
Matches 401; Conservative 180; Mismatches 308; Indels 111; Gaps 27;

Qy 8 KSNIOAGLNTSKSL-----KNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
Db 21 KSAKSGAGALKNGLGQVKQAGQKLLIYPKDYQASTGSLNDLVKAAEALGIEVHRSEK 80
Qy 61 NHTETAKSVDTVNOFLSLTQTGIAISATKLEKFLQKHSNKLAKGL-DSVENIDRKLK 119
Db 81 NGTALAKELFGTTEKLLGFSERGIALFAPQDFKLKLNK--QKLSKSLGSSSEALGORL 138
Qy 120 ASNVLSLTSFLGTALAGIELDSLKK-----GDAAPDALAKASIDLINEIIGNLSOST 175
Db 139 TQTALSAQLGFLGTATAGMDLSLRRRNGEDVSGSELAKAGVDLAAQLVDMNIASATG 198
Qy 176 IEAFSSQLAKLGTISQAKGFSNIGNKLQNL-NFSKTNLGLIITIGLISGISAGFALADK 234
Db 199 VDAPAEQGLKGNALNTR-LSGLASKLNNLPDLSLAGPGFVDAVSGILSVVSASFILSNK 257
Qy 235 NASTGKKVAGFELSNOVGNVTIAISSYLAQVAAGLSTTGVAALITSSIMLAISPL 294
Db 258 DADAGTAAAGIEISTKILGNIGKAVSQYIIAQRVAAGLSTTAATGGLIGSVVVALAISPL 317
Qy 295 AFNNAADKFNHANALDEPAKQFKFGYDGHLLAEYQGVGTIEASLTITISTALGAVSAG 354
Db 318 SFLNVADKFERAKOLEQYSEKFKFGYDGSLLASFYRETGAIEAALITINSVLSAASAG 377
Qy 355 VSAAVGSAGVGTPIALLVAGTGLISGILEASKQAMFESVANRLQKILEWEKQNGQNY 414

```


Db 378 VCAATSLVGPAAVLAISALTSGLDASKQAFERVAATKLANDEKXH-CKNY 436
Qy 415 FDGVDSYAYLANLKLSELNKELEBAERVAITQORWNNNGELAGITKLERIKSG 474
Db 437 FENGIDARSHAFLEDTFELLQYNKEYSVERVVAITQORWNNNGELAGITKSG 496
Qy 475 KAYADAFEDGKVEAGSNITLDAKT-----GLIDSNNGKTKQALHETSPLTAGTESR 529
Db 497 KAYVDFEGLKEKPEP-RFDKVKDPLEGKIDLSSIN--KTLLEKRVTVFVAGEIR 553
Qy 530 ERLTNGKYSYINKLFRKXQWQVTDGEASSKL-DPSKVIQORVAETEGTBEIGLIVAKA 588
Db 554 EKQTKGYEYTELTVKGEKXWVTCVQSHAIYDYNLIQLAIDKKG-EKQVYIESHL 612
Qy 589 G--NDDIFVQCKMINDGGDHRVFSK-DGGFGNITVDGTSATEAGSYTVNRKV-ARG 644
Db 613 GEKNDRIYLSGSSIVYAGNGHDVAYDKTDGY--LTFDQSAQKAGEYIVTKELKADV 670
Qy 645 DIYHEVVKROETKVKRTTETQYRYELR--KVGYGYOSTNLSKVERVIGSQFNDVPKG 702
Db 671 KVLKEVVKTDQISVGRSEKLEYRDELSPFELGNGIRAKDELHVSVEIIGSNRKKPFG 730
Qy 703 SKFNDFHSGEGD-----DLDGGAGDDRLFGKGNDRLSGDEGDDL 744
Db 731 SRFTDFHAGAKDDRIYGDNDHILYGGDNDVHGGDNDHVLGGNGDRLIGKGNPF 790
Qy 745 LDGSGDP-----VLNGAGNDVY-----IFRKGNDNTLYDGTGND----- 781
Db 791 LGGDGDDELQVFEQYVLLGGAGNDILYSGDGTNLPFGVGNDKIYGLGKDIYRYSK 850
Qy 782 -----KLAFADANI-----SPIMIERKEGIIVKENDHSG-SINIP 816
Db 851 EYGRHIIIEKGGDDTLTLLSLDFKDVGFIRIGDILLVKNKRIIGTLYYHEDYNGNALTIK 910
Qy 817 RWTITSNLQYOSNKTDRHKEIOLIGKSGSYITSQIDKILQDKKGTVITSQELKLADE 876
Db 911 DWF--KEGKEGQNN--KIEKIDVKGAYLSQVLTETAPGRGINFYNGLEEKLYYGE 964
Qy 877 NKSQKLSASDIASSLNKLVGSMALF-GTANSVSSNALQPI 915
Db 965 GYN---ALPQLRKDEIQIISSTGAFTGDHGVKSVSGGGL 1001

RESULT 33

US-08-258-188-2
Sequence 2, Application US/08258188
Patent No. 5475098

GENERAL INFORMATION:

APPLICANT: HALL, Robert H.
APPLICANT: XU, Jian Guo
TITLE OF INVENTION: A NEW AND DISTINCTIVE DNA SEQUENCE OF E.
TITLE OF INVENTION: coli 0157:H7 AND ITS USE FOR THE RAPID, SENSITIVE AND
TITLE OF INVENTION: SPECIFIC DETECTION OF O157:H7 AND OTHER ENTEROHEMORRHAGIC
TITLE OF INVENTION: E. coli
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/258,188
FILING DATE: 14-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DOCKET NUMBER: 15280206, DHHSEL35940
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 758 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-258-188-2

Query Match 29.6%; Score 1376.5; DB 1; Length 758;

Best Local Similarity 40.5%; Pred. No. 3.2e-99;
Matches 304; Conservative 129; Mismatches 225; Indels 93; Gaps 16;

Qy 229 FALADKNASTGKVAAGFELSNOVIGNVTKAIISSVLAQORVAAGLSTTGAVALITSSIM 288
Db 2 FILGNSDHTGTAAAGIELTQVLGNVKAQVQYILAQORVAAGLSTTGAVALITSSIM 61
Qy 289 LAISPLAFMNAADKFNHANALDEFKQFRKFGYDGHLLAEYQKRGVGTIBASLTITSTAL 348
Db 62 LAISPLSFLAADKFERAKQLESYERFKLNYEGDALLAAAFHKTGAIDAALTTINTVL 121
Qy 349 GAYSAGYSAAGVSAVGPPIALLVAGVTGLISGLEASKOAMPESVANRRLOGKILEWEKQ 408
Db 122 SSVSAGYSAASSASLIGAPISMLVSALTGTISGLEASKOAMFEHVAEKAFAARINEWEKE 181
Qy 409 NGQNYQFDKGYDRYAAYLANLKFISELNKELEAERVAITQORWNNNGELAGITKLG 468
Db 182 H-GKVFYENGDAHAAFLSDLSLLADFQRHAVERAVAITQOHWEKIGELAGITRNA 240
Qy 469 ERIKSGKAYADAFEDGKKEVAG---SNITLDKATGIIDISNNGKKTQALHFTSPLLTA 524
Db 241 DRSGQKAYINYLENGGLLEAQKPEFTQVDFPQKGTIDL--STGNVSSVLTFTTFTPT 298
Qy 525 GTSERRLTNGKYSYINKLFRKXQWQVTDGEASSK--LDPSKVIQORVAETEGTDEIGL 582
Db 299 GEEVREKQSGKYEYMTSLIVNGKDTWSV-KGIKNHGVYDYSKLIQFVKNNKHQYARI 357
Qy 583 IVNAKAGNDDIFVQCKMINDGGDHRVFSKDGKGFNITVDGTSATEAGSYTVNRKVA 642
Db 358 ISELGDKDDVYSGAGSSEVPAGEGYDVSYNKT-DVGKLTIDATGASKPEYIVSKNM- 415
Qy 643 RGD--IYHEVVKROETKVKRTTETIQRDYELKRVGYGYOSTNLSKVEEVIGSQFNDVP 700
Db 416 YGDKVLQEVVVEQESVSGKRTETIKIQRDFEFRTGGIPYDVIDNLHLSVEELIGKHDEF 475
Qy 701 KGSFNDIFHSG-----EGDLDGGAGDDRLFGKGNDRLSGDEGD 742
Db 476 KGGKFNDFHAGDNDYIEGNYGNDRLYGGDGGDYSGGQDQLFGSGNDKLSGGDGN 535
Qy 743 -----DLDGGSGDDVVLNGAGNDVYIFR 766
Db 536 NYLTGSGNDELQAHGAYNILSGTGDGDKLYGGGIDLLGGEGNDYINGFGNDIYVG 595
Qy 767 KGDGNDTLYD-GTGNDKLAFADANISIMIERKEGIIVK-----NDHSGSIN 814
Db 596 QNYGHHTIADGEGKGRDLHLSDISFDDIAFRKRVGNDLIMNAKINGVLSFNESDVNG-IT 654
Qy 815 IPRWYITSNLQYOSNKTDRHKEIOLIGKSGSYITSQIDKILQDKKGTVITSQELKLA 874
Db 655 FKWFAKD-----ASGADNHLVEVITDKGREI---KVDKIPHNNE-----RSYIYKASN 702
Qy 875 DENKSQKLSASDIASSLNKLVGSMALFETAN 905
Db 703 IASEKWNVITSVANDINKIISSVSGFDSGD 733

RESULT 34

US-08-526-813-2

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:12:24 ; Search time 39 Seconds
(without alignments)
4976.856 Million cell updates/sec

Title: US-10-069-799-5

Perfect score: 4647

Sequence: 1 MSNINVIKSNIQAGLNSTKS.....SSNALQPIPTQGTILAPSV 927

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|---------------------|
| 1 | 4642 | 99.9 | 927 | 11 | US-09-884-696-2 |
| 2 | 2325 | 50.0 | 953 | 11 | US-09-884-696-3 |
| 3 | 2326.5 | 48.1 | 956 | 11 | US-09-884-696-4 |
| 4 | 1845 | 39.7 | 1023 | 11 | US-09-884-696-5 |
| 5 | 901 | 19.4 | 695 | 11 | US-09-305-924-13 |
| 6 | 408.5 | 8.8 | 1115 | 15 | US-10-193-950A-2 |
| 7 | 295.5 | 6.4 | 1017 | 12 | US-10-369-493-20166 |
| 8 | 263.5 | 5.7 | 4327 | 12 | US-10-369-493-10178 |
| 9 | 232 | 5.0 | 1778 | 12 | US-10-238-075-749 |
| 10 | 230.5 | 5.0 | 499 | 12 | US-10-369-493-20175 |
| 11 | 226.5 | 4.9 | 2122 | 10 | US-09-813-214A-9 |
| 12 | 219 | 4.7 | 2285 | 10 | US-09-932-183A-2 |
| 13 | 216 | 4.6 | 2039 | 15 | US-10-192-584-7 |
| 14 | 211.5 | 4.6 | 2042 | 15 | US-10-192-584-6 |
| 15 | 202 | 4.3 | 2353 | 10 | US-09-797-862-33 |

| | | | | | | |
|----|-------|-----|------|----|---------------------|--------------------|
| 16 | 200 | 4.3 | 239 | 12 | US-10-369-493-19032 | Sequence 19032, A |
| 17 | 199 | 4.3 | 3241 | 9 | US-09-841-786-1 | Sequence 1, Appli |
| 18 | 197 | 4.2 | 1833 | 12 | US-10-175-275-4 | Sequence 4, Appli |
| 19 | 197 | 4.2 | 1833 | 12 | US-10-175-282-4 | Sequence 4, Appli |
| 20 | 197 | 4.2 | 1992 | 12 | US-10-175-275-3 | Sequence 3, Appli |
| 21 | 197 | 4.2 | 1992 | 12 | US-10-175-282-3 | Sequence 3, Appli |
| 22 | 196 | 4.2 | 2086 | 9 | US-09-815-242-5639 | Sequence 5639, Ap |
| 23 | 196 | 4.2 | 5795 | 9 | US-09-815-242-12610 | Sequence 12610, A |
| 24 | 191 | 4.1 | 262 | 12 | US-10-369-493-19792 | Sequence 19792, A |
| 25 | 188.5 | 4.1 | 1098 | 10 | US-09-797-862-32 | Sequence 12, Appli |
| 26 | 188 | 4.0 | 1626 | 15 | US-10-185-990-11 | Sequence 11, Appli |
| 27 | 186.5 | 4.0 | 2659 | 12 | US-10-311-879-28 | Sequence 28, Appli |
| 28 | 186.5 | 4.0 | 3169 | 15 | US-10-114-170-257 | Sequence 257, App |
| 29 | 186.5 | 4.0 | 6281 | 9 | US-09-815-242-12996 | Sequence 12996, A |
| 30 | 180.5 | 3.9 | 1136 | 12 | US-10-369-493-19046 | Sequence 19046, A |
| 31 | 180 | 3.9 | 210 | 12 | US-10-369-493-20095 | Sequence 20095, A |
| 32 | 180 | 3.9 | 273 | 12 | US-10-369-493-20096 | Sequence 20096, A |
| 33 | 179 | 3.9 | 773 | 9 | US-09-841-786-6 | Sequence 6, Appli |
| 34 | 179 | 3.9 | 992 | 12 | US-10-193-764-57 | Sequence 57, Appli |
| 35 | 179 | 3.9 | 998 | 12 | US-10-193-764-55 | Sequence 55, Appli |
| 36 | 177.5 | 3.8 | 2434 | 9 | US-09-815-242-5835 | Sequence 5835, Ap |
| 37 | 176.5 | 3.8 | 1741 | 10 | US-09-971-536-68 | Sequence 68, Appli |
| 38 | 175.5 | 3.8 | 1228 | 10 | US-09-117-447-2 | Sequence 2, Appli |
| 39 | 175 | 3.8 | 2893 | 12 | US-09-882-227-522 | Sequence 522, App |
| 40 | 174.5 | 3.8 | 1536 | 12 | US-10-193-764-63 | Sequence 63, Appli |
| 41 | 174 | 3.7 | 2124 | 12 | US-10-369-493-11841 | Sequence 11841, A |
| 42 | 172 | 3.7 | 637 | 12 | US-10-369-493-14403 | Sequence 14403, A |
| 43 | 172 | 3.7 | 2834 | 12 | US-10-085-959-252 | Sequence 252, App |
| 44 | 170.5 | 3.7 | 1536 | 14 | US-10-092-880-2 | Sequence 2, Appli |
| 45 | 169 | 3.6 | 1073 | 12 | US-10-193-764-45 | Sequence 45, Appli |

ALIGNMENTS

RESULT 1

US-09-884-696-2
; Sequence 2, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
; APPLICANT: GEORGE, LISLE W
; APPLICANT: ANGELOS, JOHN A
; APPLICANT: HESS, JOHN F
; TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
; TITLE OF INVENTION: AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
; TITLE OF INVENTION: BOVIS INFECTIONS
; FILE REFERENCE: 481.06
; CURRENT APPLICATION NUMBER: US/09/884,696
; CURRENT FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 927
; ORGANISM: Moraxella bovis
US-09-884-696-2

| | | | | | | | |
|-----------------------|--------|---|-------|------------|-----|--------|------|
| Query Match | 99.9%; | Score | 4642; | DB | 11; | Length | 927; |
| Best Local Similarity | 99.9%; | Pred. No. | 0; | | | | |
| Mismatches | 926; | Conservative | 0; | Mismatches | 1; | Indels | 0; |
| Gaps | 0; | | | | | | |
| Qy | 1 | MSNINVIKSNIQAGLNSTKSLGNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAREP | 60 | | | | |
| Db | 1 | MSNINVIKSNIQAGLNSTKSLGNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAREP | 60 | | | | |
| Qy | 61 | NHTETAKSVDTNQFSLTQTGTIAISATKLEFLQKHSNKLAKGLDSVDENIDRKLGA | 120 | | | | |
| Db | 61 | NHTETAKSVDTNQFSLTQTGTIAISATKLEFLQKHSNKLAKGLDSVDENIDRKLGA | 120 | | | | |
| Qy | 121 | SNVLSTLSSFLGTALAGIELDSLIIKGDAAADALAKASIDLINEIGNLSQSTQTIKAFS | 180 | | | | |
| Db | 121 | SNVLSTLSSFLGTALAGIELDSLIIKGDAAADALAKASIDLINEIGNLSQSTQTIKAFS | 180 | | | | |

191 SOLAKLGSTISQAKGFSNIGNKLNQNFNKTNLGLIITGLSGISAGFALADKNASTGK 240
191 SOLAKLGSTISQAKGFSNIGNKLNQNFNKTNLGLIITGLSGISAGFALADKNASTGK 240
241 KVAAGFELSNOVIGNVTKAISSYVLAQVAAGLSTTGAAALITSSIMLAISPLAFMNA 300
241 KVAAGFELSNOVIGNVTKAISSYVLAQVAAGLSTTGAAALITSSIMLAISPLAFMNA 300
301 DKFNANALDEFKQFRFGYDGHLLAAYQGVGTIEASITTTISTALGAVSAGVSA 360
301 DKFNANALDEFKQFRFGYDGHLLAAYQGVGTIEASITTTISTALGAVSAGVSA 360
361 GSACTPTALLVAGVTGLISGLEASKOAMFESVANRLOGKLEWEKONGGONFEDKGYD 420
361 GSACTPTALLVAGVTGLISGLEASKOAMFESVANRLOGKLEWEKONGGONFEDKGYD 420
421 SRYAAYLANNKFLSELNKELEAERVAITQORWNNIGELAGITKGERIKSGKAYADA 480
421 SRYAAYLANNKFLSELNKELEAERVAITQORWNNIGELAGITKGERIKSGKAYADA 480
481 FEDGKVVAGSNITLDKATGIIIDISNSNGKKTQALHFTSPLLTAGTESRRLTNGKYSYI 540
481 FEDGKVVAGSNITLDKATGIIIDISNSNGKKTQALHFTSPLLTAGTESRRLTNGKYSYI 540
541 NKLFGRVKNQVTDGEASSKLDKFSKVIQVAETEGTDEIGLIVNAKAGNDDIFVGGK 600
541 NKLFGRVKNQVTDGEASSKLDKFSKVIQVAETEGTDEIGLIVNAKAGNDDIFVGGK 600
601 NIDGDDGHRVYFYSKGGFGNITVDGTSATGAGSYTVNRKVARGDIYHEVVKRQETK 660
601 NIDGDDGHRVYFYSKGGFGNITVDGTSATGAGSYTVNRKVARGDIYHEVVKRQETK 660
661 RTETIQRDYELRVKGYQSTDNLSKVEEVIGSQFNDVFKSGFNDIFHSGEGDLDG 720
661 RTETIQRDYELRVKGYQSTDNLSKVEEVIGSQFNDVFKSGFNDIFHSGEGDLDG 720
721 GAGDRLFGGKGNDRLSGDEGDDLDGSGDDVNLGGAGNDVYFRKGNDTLVDGNG 780
721 GAGDRLFGGKGNDRLSGDEGDDLDGSGDDVNLGGAGNDVYFRKGNDTLVDGNG 780
781 DKLAFADANISDIMEIKTEGIIIVKRNHSGSINIPRWYITSNLQYQSNKTDHKEOLI 840
781 DKLAFADANISDIMEIKTEGIIIVKRNHSGSINIPRWYITSNLQYQSNKTDHKEOLI 840
841 KGDSYITSDQDKILOKDGTVITSOELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
841 KGDSYITSDQDKILOKDGTVITSOELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
901 FGTANSVSSNALOPTQTOGILAPSV 927
901 FGTANSVSSNALOPTQTOGILAPSV 927

RESULT 2

US-09-884-696-3

; Sequence 3, Application US/09884696

; Publication No. US20030035809A1

; GENERAL INFORMATION:

; APPLICANT: GEORGE, LISLE W

; APPLICANT: ANGELOS, JOHN A

; APPLICANT: HESS, JOHN F

; TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES

; TITLE OF INVENTION: AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA

; TITLE OF INVENTION: BOVIS INFECTIONS

; FILE REFERENCE: 481.06

; CURRENT APPLICATION NUMBER: US/09/884,696

; CURRENT FILING DATE: 2001-06-19

; NUMBER OF SEQ ID NOS: 41

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 3

; LENGTH: 953

; TYPE: PRT

; ORGANISM: Pasteurella haemolytica

US-09-884-696-3

Query Match 50.0%; Score 2325; DB 11; Length 953;

Best Local Similarity 49.9%; Pred. No. 2.5e-158;

Matches 464; Conservative 175; Mismatches 264; Indels 26; Gaps 13;

QY 8 KSNIQAGINSTKSLGNLYLAIPKD--YDPQGGTINDFIKAADDELGIARLAERPNHPT 65
DB 28 QSLTQAG--SSUKTAKKIIIIYIPQYDTEQNGQLQDLVKAABELGLEVQREERNIAT 86
QY 66 AKKSVDTVNPFSLTQTGTIAISATKLEKFLQKHSNKLAKGLDSVENIDRKLKASNYLS 125
DB 87 AOTSIGTIQTAIGTERGIVLSAPQIDKLQX---TRAGQALGSAESIVQANKAKTVLS 143
QY 126 TLSFELGTALAGIELDSLIKKGDAAPDALAKASIDLNEIIGLSQSTQTTEAFSSQLAK 185
DB 144 GIOSILGSVLGMDLDEAL--QNNNOHALAKAGLELNSLIENTANSYKTLDFEGEQISQ 202
QY 186 LGSTISQAKGFSNIGNKLNQNFNKTNLGLIITGLSGISAGFALADKNASTGKVVAA 244
DB 203 FGSKLQNIKGLCTUGDKNIGGLDKAGLDVLSGLLSGATAALVLADKNASTAKKVA 262
QY 245 GFELSNQVIGNVTKAISSYVLAQVAAGLSTTGAAALITSSIMLAISPLAFMAADKFN 304
DB 263 GFELANQVGNITKAVSSYILAQVAAGLSTGPAALIASTVSLAISPLAFAGIADKFN 322
QY 305 HANALDEFKQFRFGYDGHLLAAYQGVGTIEASITTTISTALGAVSAGVSA 364
DB 323 HAKLSYAEFRFKLGYDGNLLAAYQGVGTIDASVTAINATAALAGGYSAAAGSVI 382
QY 365 GTPITALLVAGVTGLISGLEASKOAMFESVANRLOGKLEWEKONGGONYPDKGDSRYA 424
DB 383 ASPITALLVAGVTGLISGLEASKOAMFESVANRLOGKLEWEKONGGONYPDKGDSRYA 442
QY 425 AYLANNLKFLSELNKELEAERVAITQORWNNIGELAGITKGERIKSGKAYADAPEDG 484
DB 443 ANLQNMKFLNLNKELEAERVAITQORWNNIGELAGITKGERIKSGKAYADAPEDG 502
QY 485 KKVAGSNITLDKATGIIIDISNSNGKKTQALHFTSPLLTAGTESRRLTNGKYSYINKL 544
DB 503 KHIKADKLVLQDSANGIIDVSNKAKTQHIPLFTPLTTPGTEHRRVQTKYIITKLN 562
QY 545 FGRVKNQVTDGEASSKLDKFSKVIQV---AETEGTDEIGLIVNAKAGNDDIFVGG 597
DB 563 INRVDSWKIIDGAASSTFDLTNVVQRIEGLDNAGNVTKETKIIAKLGEGDNNVFGS 622
QY 598 GOMNIDGGDGHDRVYFYSKGGFGNITVDGTSATGAGSYTVNRKVARGDIYHEVVKRQETK 657
DB 623 GTTEIDGGEGYDRVHYSR--GNYGALTIDATKETEGQSYTVNRVFTETGKALHEVTSHTAL 681
QY 658 VGRTEITQYRDYELRVKGYQSTDNLSKVEEVIGSQFNDVFKSGFNDIFHSGEGDLD 717
DB 682 VGNREEKIEYR--HSNNQHHAGYITKTLKAVEEIIIGTSHNDIFPKSGFNDAFNGDGVDT 740
QY 718 LDGAGDRLFGGKGNDRLSGDEGDDLDGSGDDVNLGGAGNDVYFRKGNDTLVDG 777
DB 741 IYGNNDGNDRLFGGKGGDDI LDGNGDDDFIDGGKNDLHGGKGGDDI FVHRKGGDNDIIDS 800
QY 778 TGNDKLAFADANISDIMEIKTEGIIIVKRNHSGSINIPRWYITSNLQYQSNKTD 833
DB 801 DGNDKLSFSDSNLKDITFEKVKHNLVI--TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 858
QY 834 HKIEQLTGKDSYITSDQDKILOKDGTVITSOELKKLADENKSQKLSASDIASSLNK 893
DB 859 EKIEEIIQNGERITSKQVDDLI--AKNGKIKITODELSKVVDNYELLKHS--KNVTNSLDK 915
QY 894 LVGSMALFGTANSVSSNALOPTQTOGI 922
DB 916 LISSVSAFTSSNDRNVIVAPTSMLDQSL 944

RESULT 3

US-09-884-696-4

```
; Sequence 4, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
; APPLICANT: GEORGE, LISLE W
; APPLICANT: ANGELOS, JOHN A
; APPLICANT: HESS, JOHN F
; TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
; TITLE OF INVENTION: AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
; TITLE OF INVENTION: BOVIS INFECTIONS
; FILE REFERENCE: 481.06
; CURRENT APPLICATION NUMBER: US/09/884,696
; CURRENT FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 956
; TYPE: PRT
; ORGANISM: Actinobacillus pleuropneumoniae
US-09-884-696-4

Query Match      48.1%; Score 2236.5; DB 11; Length 956;
Best Local Similarity 47.7%; Pred. No. 5.8e-152;
Matches 453; Conservative 188; Mismatches 255; Indels 53; Gaps 15;

Qy      1 MSNTNV--IKSNIQAGL-----NSTKGLKNLYLAIPKDYDPQ 36
Db      1 MSKITLSLKSLSLOGLKNGKNLNQAGTTLKNGLTQTGHSLQNGAKKLIYIPQYDSG 60

Qy      37 KGGTLDNFIKAADELGIARLAEENPHHTETAKSVYDVTNQFSLTQTGTGIAISATKLEFLQ 96
Db      61 QGNGVQDLVKAANDLGIWREERSNLDIAKTSFDTTQKILGFTDRGIVLFPAPQLDNLK 120

Qy      97 KHSTNKLAKGLDSVENDRKLKGNASVLTSLSSFLGTALAGIELDSLKKGDAAPDALAK 156
Db      121 KNP--KIGNTLGSASSISONTKANTVGGIQTSLGSLVGLNELLKNDPNQLELAK 178

Qy      157 ASIDLINEIINLSQSTQTIEAFSSQLAKLSTISQAQFSGNIGNKLNQL--NFSKTNLGL 215
Db      179 AGLBTLWELVGNIASSVQTVDAFAEQISKLGSHLQNVKGLGLSNKLNQLNLDLQKASLGL 238

Qy      216 EITLGLSGLSAGFALADKNASTGKVAAGFELSNOVIGNVTKAISSYVLAQRVAAGLST 275
Db      239 DIISGLLSGASGLILADKEASTEKKAAGVEFANQIIGNVTKAVSSYILAQRVASGLSS 298

Qy      276 TGAAVAALITSIMLAISPLAFMAADKENHANALDEFAKQPKRGYDGDHLLAERYQGVG 335
Db      299 TGPVAALITASTVALAVSPLSFLANVADPKQADLIKSYSERFQKLGYPGDRLADPHRETG 358

Qy      336 TIEASLTITSTALGAVSAGVSAAGVSAVGPPIALLVAGVTGLTSGILEASKQAMFESVA 395
Db      359 TIDASVTITNTALAAISGGVGAASAGSLVGPVALLVAGVTGLTITILEYSKQAMFEHVA 418

Qy      396 NRGKILEWEKONGQONYPDKYDSRYAAYLANNLKFLSELNKELEAERVAIATQQRWD 455
Db      419 NKVHDRVIEWEKKH--NKNYPEQGYDSRHLADLQDMKFLINLKELOAERVAITQQRWD 477

Qy      456 NNIGELAGITKLGRIKSGKAYADAPEDGKVEKAGSNITLDAKGIIDISNSGKKTQAL 515
Db      478 NQIGDLAAISRRTDKISSGKAYVDAFEQGYQHQSDSSVQLDNKNGIINISNTN--RKTQSV 536

Qy      516 HFTSPLTATGPESERLTNGKYSYINKLKFRGVKNQWVTDCEASSKLDFFSKVIRVA--- 572
Db      537 LFRTPLTTPGEENRERQEGKNSYITKLHIQRVDSWVTVDGASSVDFTNVVQRIAVKF 596

Qy      573 -----ETEGTDBIGLTVNAKAGNDIDFVQCGKMNIDGGDHRVFSKDGFGNITVD 625
Db      597 DDAGNIIESKDTK---IIANLAGNDNVFVGSSTTVIDGGDHRVHYSR--GEYGAIVD 652

Qy      626 GTSATEAGSVTVNRKVARGDYIHEVVKQETKVKRRTETIORYDELKVKGYGQSTDNL 685
Db      653 ATAETEGSVKRYVGVGSKALHETIATHQTVNGNREEKIEYR--REDDRPHGTGYTVDLS 711

Qy      686 KSVEEIVGSQFNDVPKGSKFNDIFHSGEGDLDLGGAGDRLFGCKGNDRLSGDEGDDL 745

; Sequence 5, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
; APPLICANT: GEORGE, LISLE W
; APPLICANT: ANGELOS, JOHN A
; APPLICANT: HESS, JOHN F
; TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
; TITLE OF INVENTION: AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
; TITLE OF INVENTION: BOVIS INFECTIONS
; FILE REFERENCE: 481.06
; CURRENT APPLICATION NUMBER: US/09/884,696
; CURRENT FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-884-696-5

Query Match      39.7%; Score 1845; DB 11; Length 1023;
Best Local Similarity 41.8%; Pred. No. 9.2e-124;
Matches 421; Conservative 156; Mismatches 311; Indels 118; Gaps 23;

Qy      8 KSNIQAGLNSTKGLKNLYLAIPKDYDPQKGGTLDNFIKAADELGIARLAEENPHHTETAK 67
Db      32 KDALKKAAEOTRAGNELILLIPKDYKGO--GSSLNDLVRTADELGIQVQYDEKNGTAITK 90

Qy      68 KSVDTVNFQSLTQTGTGIAISATKLEKFLQKH--STNKLAKGLDSVENIDRKLKGNASVLS 125
Db      91 QVFGTAEKLGITGERGVTFAPQLDKLLQYKAGNKLK---GSAENIGNLGRAGSVLS 147

Qy      126 TLSPLGTALAGIELDSLKK---GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSS 181
Db      148 TFQNFGLTALUSSMKIDELIKKQSGGNVSSSELAKASIELINQLV--DTAASLNNVNFSQ 206

Qy      182 QLAKLGSTISQAQFSGNIGNKLNQL--NFSKTNLGLTITGLSGISAGFALADKNASTGK 240
Db      207 QLNKLSVLTNTHLNGVGNKLNQLNPLNDNIGAGLDTVSGILSAISAFILSNADATGT 266

Qy      241 KVAAGFELSNOVIGNVTKAISSYVLAQRVAAGLSTTGAAVAALITSSIMLAISPLAFMAA 300
Db      267 KAAAGVELTTKVLGNVGKISQVYIAQRAAQGLSTSAAGLIASVVVTLAISPLSFLSIA 326

Qy      301 DKFNHANALDEFAKQPKRGYDGDHLLAERYQGVGTGTEASLITITSTALGAVSAGVSAAV 360
Db      327 DKFRANKIEYSQRFKGLGYDGDLSLAAAFHKETGADASLSTRISTVLAIVSSGISAAT 386

Qy      361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLOQKILEWEKONGQONFYDKGYD 420
Db      387 TSLVGAPVSLVAGVTGIIISGILEASKQAMFEHVASKMADVIAEWKKH--GKNYFENGVD 445

Qy      421 SRVAAAYLANNLKFLSELNKELEAERVAIATQQRWDNNIGELAGITKLGRIKSKAYADA 480
Db      446 ARHAAFLDKNFKILSQYNKEYSVERSVLIITQHQWDTLIGLAGVTRNGDKTSLGSKSYIDY 505
```



```
Qy 481 PEDCKKVEAG-----SNITLDAKTGIIDISNSNGKTKQALHFTSPLLTAGTESRRLTNGK 536
Db 506 YEEGRLEKKPDEQKQVDFPKGNIDLSDS--KSSTLLKFEVTPLLTPGEEIRERROSGK 563
Qy 537 YSYINKLKFGKVKWQVTD--CEASSKLDPSKVIQORVAETECT--DEIGLIVNAKAGNDIF 594
Db 564 YEYITELLVKGVDRKWTGVQDQKGSVDYNSNLIQHASVGNNOYREIRIESHLGGDDKVP 623
Qy 595 VGOQKNIIDGGDHRDVPYSK--DGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVV 651
Db 624 LSAGSANIYACKGHVYVYDXTDGY--LTIDGTEAGNATVTR--VLGGDVKVLQEVV 680
Qy 652 KRQETVKGRTRTTOYRDYELRKV--GYGYQSTDNLKSVEEVIGSQFNDVFKGSFNDIFH 710
Db 681 KEQSVSGKRTKTEQYRSYEFTHINGKMLTETDNLYSVEELIGTTRADKPFKGSFADIFH 740
Qy 711 SGEGLDLDGAGDRLF----- 728
Db 741 GADGDHIEGNDGNDRLYDGNKNTLSCGNGDDQLYGGDGNDKLIGGAGNNYLNCGDGD 800
Qy 729 -----GGKGNDRLSGDEGDDLLDGGSGDDVLNGAGNDVYIFRKGDNNDT 773
Db 801 ELQVQNSLAKXNVLSGGKGNKLYGSEGADLLDGGEGNDLLKGGVGNDIYRLSGYGH 860
Qy 774 LYDTGT--NDKLAFRADANISDIMEKTEGIIIVKRD-----HSGSINIPRWITSNLQ 825
Db 861 IDDDGGKDDKLSADIDPRDVAFRREGNDLIMYKAEGNVLSIGHKNGITFKXNF-----E 915
Qy 826 NYQSNKTDHKEQLIGKDGSVITSDQIDKILQDKK-----DGTVITSQ----- 868
Db 916 KESGDISHQIEQFDKDRVITPDSLKKALEYQOSNNKASYVYVGNDAALYGSQGNLPL 975
Qy 869 --ELKKADENKSOXLSADSIASSINKLVGSMALFGTA--NSVSSNA 911
Db 976 INEISKIISAGNFVDKVEERAASLLQLSGNASDFSYGRNSITLTA 1021

RESULT 5
US-09-305-924-13
; Sequence 13, Application US/09305924A
; Publication No. US20030091579A1
; GENERAL INFORMATION:
; APPLICANT: Jack G. Manns
; APPLICANT: Stephen D. Acres
; APPLICANT: Richard Harland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,217
; EARLIER FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 13
; LENGTH: 695
; TYPE: PRT
; ORGANISM: GnRH
US-09-305-924-13

Query Match 19.4%; Score 901; DB 11; Length 695;
Best Local Similarity 38.8%; Pred. No. 3.9e-56;
Matches 221; Conservative 99; Mismatches 165; Indels 84; Gaps 14;

Qy 14 GLNSTKSLGNLYLAIPKD--YDPQKGGTLDNDFIKAADELGIARLAEEPNTETAKSV 71
Db 109 GSSPPTGAKKILYIPQNYQYDTQGNQLQDLVKAEEELGIEVQREERNIATAQTSLG 168
Qy 72 TVNQPLSTQGTGIAISATKLEFKLOKHTNKLAKGLDSVENIDRKLKASNVLSLSSFL 131
Db 169 TIQTAGITGERGIVLSAPQIDKLOK---TRAGOALGSAESITVQNKAKTVLSGIQIL 225
Qy 132 GTALAGIELDSLKKGAAPDALAKASIDLINEIGNLSQSTQTTEAFSSQLAKLGSTIS 191
```

```
Db 226 GSVLWAGMDLEAL--QNNSNQHAKAGLELTNSLIENIANSVKTLDFEQISQFGSKLQ 284
Qy 192 QAKGFSNIGNKQNL--NFSKTNLGLEIITGLLGSISAGFALADKNASTGKKVAAGFELSN 250
Db 285 NIKGLGTGLGDKLJNIGGLDRAGLGLDVISGLSGATAALVLADKNASTAKKVGAGFELAN 344
Qy 251 QVTGNVTKATSSVLAQORVAAGISTTGAVAALITSSIMLAISPLAFNNAADKFNHANALD 310
Db 345 QVGNITKAVSSYILQORVAAGSSTGPFVAALITASTVLSAISPLAFAGIADKFNHAKSLE 404
Qy 311 EFQKFRKFGYDGDHLLAEYQRGVGTIEASUTITSTALGAVSAGVSAAGVSGVGTPIAL 370
Db 405 SYAERFKLGYDGDNLAEYQRGVTGIDASVTALNTALAAIAGGVSAAA----- 453
Qy 371 LVAGVTGLISGLEASKQAMFESVANRLOSKLILEWKQNGQNYFDKGYDSRYAAYLANN 430
Db 454 --ADLT-----FEKVGNLVTNSKKKXVT--IONWF----- 481
Qy 431 LKFLSELNKELEAERVIAITQQRWDNNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAG 490
Db 482 -----READFAKEVFNKYKATKDEKIEETIG--QNGERITTSKQV-----DGLIAGN 525
Qy 491 SNITLDAKTGIID-----ISNSNGKKTQALHFTSPLLTAGTESRRL-----TNGKY 537
Db 526 GKITQDELSKVVDNYELLKHSKNVTNSLDKLISSV-----SAFTSSNDSRNVLVAPTSM 581
Qy 538 SYINKLKFR--VKWQVTDGEASKLDFS 565
Db 582 QSLSSLQFARGSQHWSYGLRPGSGSQDWS 610

RESULT 6
US-10-193-950A-2
; Sequence 2, Application US/10193950A
; Publication No. US20030104002A1
; GENERAL INFORMATION:
; APPLICANT: Sparling, P. Frederick
; APPLICANT: Thompson, Stuart
; TITLE OF INVENTION: Antigenic iron repressible proteins from N. meningitidis related
; FILE REFERENCE: hemolysin family of toxins
; CURRENT APPLICATION NUMBER: US/10/193,950A
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 09/045,177
; PRIOR FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 2
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Neisseria Meningitidis
US-10-193-950A-2

Query Match 8.8%; Score 408.5; DB 15; Length 1115;
Best Local Similarity 20.4%; Pred. No. 2.1e-20;
Matches 239; Conservative 171; Mismatches 384; Indels 399; Gaps 51;

Qy 40 TLNDFIKAADELGIARLAEEPNTETAKSVYDVTYNQFLSL-----TQTG----- 83
Db 55 TVNNAIQADSF--LSSIGRDNKITNTASLLASLONIPLNLRNVSRDIRETGKFKPNDIQ 113
Qy 84 ----IATSA-----TKLEKFLQ--KXSTNKLAKGLDSVENIDRKLKASNVLSLSS 130
Db 114 AIGDIFIAAGDGLQYIKQOTEAMAQSRFLPTKLTGLNDV--LNSRMLKSTVLQHELNY 171
Qy 131 LGTALA--GIEL--DSLKKGAAPDALAK--ASIDLINEIGNLSQSTQTEAFSSQLA 184
Db 172 LGFKIKYGNRIGESIMNIDDDPTPSKIANFFADPDYISNVLEVSRYISLVDPDANPW 231
Qy 185 K-----LGSTISQ----- 192
Db 232 KGGEDYIGRGISEWGELEKWKYKQDFLYLEKEWDQFPKFDWLPFEPFENAREWLKLDPK 291
```

193 QY -----AKFSNIGKLNLFKSNLGLLEIITGLSGISAG 228
 292 DB RSGKYHYVDPLALDLDGDIETVAAGFA-----GALFDRHQGGIRTAGWISA-DDG 343
 229 QY FALADKNASTGKKAAGFELSNVIGNVTKAISYVLAQRAVAGLSGTGAVALITSSIM 288
 344 DB LLVRDLN-----GNGIIDNGAELFGDNT---KLADGSPAKHGVAALAELE-- 395
 289 QY LAISPLAFMAADKFNHANALDEPAKQFRKPGVDG-----DHLAAYQORGVTTEASLTTI 344
 386 DB -----SNGDNIINAADAFAQTLRWQDLNODGISOANELRTLELIGIQLSDLAYKDV 437
 345 QY STALGAVSAGVSAAGVACTPIALLVAGVTGLISGLEASQAMPESVANRLQKILE 404
 438 DB KNLG-----NGNTLAQQSGYTKTD-----GTTAKMGDILLAA-----DNLHSPKDKVEL 483
 405 QY WEKQNGQNYFDKG--YDSRYAAYLANNLKPLSELNKELEA-ERVIAITQ---QRW---D 455
 484 DB TAEQAKAANLAGIGRLDREAAALSGDLANMLKAYSAAETKEAQLALLDNLHKWAETD 543
 456 QY NNTGELA-----GITYLGERI-----KSGKAYADAFEDGKKVEA 489
 544 DB SNMGKSPMLSTDTWTQANEGIALTPSQVAQLKKNALVSLDKAKAAIDAARDRIAV-- 601
 490 QY GSNITLDAKTG-----IIDSNSN----- 508
 602 DB LDAYTQDSSTLYMSEEDALNIVKNTDYLHLAKNIYQNLFPQRLQPYLNQI 656
 509 QY -----GKKTQALHFTSPLLTAGTESRRLTNGKYSYN---KLKFRGVKNW-----Q 552
 657 DB SFRKENDTFTLDF-SGLVQAFNHVKE--TNPQKAFVDLAEMLAAGELRSWYEGRRLMADY 713
 553 QY VTGGEASKL-DPSKVQR-----VAETGTEDEGLIVN-----AKAGNDIP 594
 714 DB VEEAKKAGKPEYOKVLQGOETVALLAKTSGTQADAILQNVFGHKNKVSLYGNDGNDTLI 773
 595 QY VGQCKMNDGCGDHRVPYSKGGFGNITV-----DGTSA-----TEAG 633
 774 DB GGAGNDYLEGGSGSDTVVFGK--GFGQDTVYNYDYATGRKDIIRFTDGTADMULTFTREG 831
 634 QY SY-TVNRKRVARGDIYHEVVRQETKVGKRTETIQYRDYELRKVG----- 676
 832 DB NLLIKAKDSSGVTVQSYFQNDGSGAYRIDEIHFNGKVLVDVATKVELVQQSDGSDRL 891
 677 QY YGYOSTNLK---SVEEVIGSQFNDVPKSGKFNDI FHSGBGDDLLDGGAGDRLF----- 728
 892 DB YAYQSGSTLGGGLGDDYLYGADGNLLNGDAGNDIYSGNGNDTLDDGEGNDALYNGN 951
 729 QY -----GKGNDRLSGDEGDDLLDGGSDVNLGAGAGNDVYI FRKGDGNDTLVD---GTGND 781
 952 DB DALMGGBGNDHLNGEDGNDTLIGGAGNDYLEGGSGSDTVYFPEGFGQDTVYNYHVDKNSD 1011
 782 QY KLAFADANI SDIMERTKEGIIKVRNDHSGSINIPRWYITSNLQYQSNKTDHKIEQLIG 841
 1012 DB TMHFKGKAADVHFRSGSDLVLSASEQ--NVRISGFFYGEN-----HRVDIFVP 1060
 842 QY KDGSIYTSDDIKLQDKKDGTVITSQELKLADEKNSQKLSASDIASSL---NKLVGSM 898
 1061 DB DDA-----TANSYSSNALQPIPTQGILAPS 926
 899 QY ALFG-----TANSYSSNALQPIPTQGILAPS 926
 1085 DB SVFGSNTAATGNGVDAN-IQSVQOQL--LVTPS 1114

RESULT 7
 US-10-369-493-20166
 ; Sequence 20166, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfang
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 20166
 ; LENGTH: 1017
 ; TYPE: PRT
 ; ORGANISM: No. US20030233675A1loc punctiforme
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(1017)
 ; OTHER INFORMATION: unsure at all Xaa locations
 US-10-369-493-20166

Query Match 6.4%; Score 295.5; DB 12; Length 1017;
 Best Local Similarity 22.6%; Pred. No. 2.5e-12;
 Matches 182; Conservative 92; Mismatches 282; Indels 249; Gaps 35;
 120 QY ASNVLTSLSSFLGTALAGIEL---DSLIIKKGDAAPDALAKASIDLINELIICNLSQSTQTI 176
 331 DB ASSNNTLDGAGANDTINAEPSGDNLLSGCD-GNDYLSSTSGTKSDNGQLNFTSEGNNTL 389
 177 QY EAFSPOLAKLGSTISQAKFSNIGNKLNQNFNFKTNLGLLEII-----TGLLSGISAGFAL 231
 390 DB DCGAGD-DRLYADISTGNLLSGCDGNDTLSTGTSTRNSFIVPSSGNNTLNG-GAGDDY 447
 232 QY ADKNASTGKKAAG-----FELSNQVIGNVTKAISSYVLAQRAVAGLSGTGAVALITS 285
 448 DB LYASASTGSNFFSGDGNDSFSLT--TISDGSFNSYLVTVQVDG-----KG 494
 286 QY SIMLATSPLAFMAADKFNHANALDEPAKQFRKFGVDGDLHLAEYQRGVGTIEASLTTI- 344
 495 DB DLLSFNLLPF-----HSDSNVA-----GGI---ITFN 521
 345 QY -STALGAVSAGVSAAGV-----SAVGTPIALLVAGVTG---LISGLEASKQAMFESVA 395
 522 DB PTINKGTITAGTVGVNYKNIKLNISGTAADDLLILSGNDTLSTG----- 567
 396 QY NRLOQKILEWKGQNGQNYED--KGYDSRYAAYLANNLKPLSELNKELEAERVIAITQOR 453
 568 DB -----NGGKDTIDGGKDDVISIDYSNATSKIITTFNATNTTGLITAGTYQI 614
 454 QY WDNNGIGELAGITKLGRIK-SGKAYADAF-----EDGKKEAGSNITLDAKTG--IIDI- 504
 615 DB SYKNI-----ERLNTGTAYDDNIIILNNGNDTLSTGSGNDTLINGGQDDILLSD 664
 505 QY --SNSNGKKTQALHFTSP-LLTAGTESRRLTNGKYSYINIKLXF-GRVKNQVTDGSS 560
 665 DB YNTNTTGITTTNTTINAGVITAGS-----NRVSYKNIERLDSGTVDYIIGNGND 718
 561 QY KLDFSKVIOQVAETGTEDEGLIVNAKAGNDIIFVGGKKNIDGGDGHDRVY-----SK 615
 719 DB TL-----SGGYDGNDDTIIGG-AGNDVWVGKGNDDILTGTVGNDKFPVYDLESNY 766
 616 QY DGG-----FGNI----- 622
 767 DB DTGTDITDFGGIKGKUSNPQAVIASLDTLEFDNSGFRFLPSAQLQLTQNGNLEITFE 826
 623 QY -----TVDGTSAEAGSYTVNRKVARGDIYHEVVRQETKVGKRTET 664
 827 DB DLFNSSNSKVLQNFQLENDGENILFPQGTITNSI---DVF-DANSTQSLINKNTVT 882
 665 QY IQYRDYELRKVGQYQSTDNLKSVEEVIGSQFNDVPKSGKFNDI FHSGBGDDLLDGGAGD 724
 883 DB F-----LNDLNNITGFDN-----SNDVWVGQGGDDIINGLSGNDLLRGNAGN 925

Qy 725 DRLFGKGNRLSDGEGDLDGSDVLLNGAGNDVYIFRKGNDTLTYDGTGNDK-- 782
Db DTLGGAGNDTLYGAGNDLTYGAGSDLLTGNDGNDIFAPAGEGTDTITDFTKNNDLI 985
Qy 783 -----LAPADANIS--DIMIERKTE 800
Db 986 GLYGGLSFGQLSFGSNNILVTSTNE 1010

RESULT 8

US-10-369-493-10178
; Sequence 10178, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 473/74
; SEQ ID NO 10178
; LENGTH: 4327
; TYPE: PRT
; ORGANISM: magnetite-containing magnetic coccus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(4327)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-10178

Query Match 5.7%; Score 263.5; DB 12; Length 4327;

Best Local Similarity 19.8%; Pred. No. 4.5e-09;
Matches 203; Conservative 123; Mismatches 358; Indels 339; Gaps 44;

Qy 1 MSNINVIKSNIOAGLN-----STKSGLNKXLYLAIPKDYDPQKGGTLNDFIKAADELGIARL 56
Db 3165 VSNITIDSGNVYANAGGATITVNDGIANLI-----TDAGTVVGTGRN--VTYDASMAQL 3218
Qy 57 AEEPNNHTAKSV---DTVNQFLSLTQGI---AISATKLEKFLQKH----- 98
Db 3219 SQIDNYTTGALKYVYIKDAVAALVANTNSYVTGSYAVSVTDAASMAQLSAIDQDTTGLT 3278
Qy 99 -----STNKLAKG-----LDSVENID-----RKLKASNVLS 125
Db 3279 YTKLTDVANLVNTNSYVTGSVNVTVSDIATISQLSSIDANTTGSVYTYTQIGDAAATLA 3338
Qy 126 T-----LSSFLGTALAGIELDSLIIKKGDAAPDALAKASIDLIN 163
Db 3339 TNAGNVKATHTVTTDAATTAQLTTIDGNTTG-----SLVYTAGGVKDSANLVVNTNS 3394
Qy 164 EIIGNLSQS--TOTIEAFSSQLAKG-----STISQAKGF--SNIGNKLONLN----- 207
Db 3395 YVTGAVNVSVTDTVSI--AQLSAVDEYTTGLTYGAGVKDSVANILLVNTNSYVTGSYAVS 3452
Qy 208 ----FSKNIGL--EIIITGLSGISAGFALADKXNASTCKKVAAGFELS---NOVIGNV-- 256
Db 3453 ITDVASMANLSAIDQFTTGTGLNLYTKLSDTVSALVANTNSYVTGSVNVTVITDNASMANNSA 3512
Qy 257 ----TKAISSVYLAQRAVAGLS-----TTGAVAALITSSIMLAISPLAFNAA--DKFN 304
Db 3513 IDONTGTLTYTKLSDTAALAANTNSYVTGSVNVTVTDN--ATVAQLTTVDAATGTGIIK 3570
Qy 305 HANALDEFAKOPRKFY--DG-----DHLLAEYQRGVGTIEASLTTI----- 344
Db 3571 VASVVDGSGNISNPFAYDGLGVSVYVANDNVMA-----VTVAQATDATVTIADDDVLT 3625

Qy 345 -----STALGAVSAGVSA-----AAVGSAGVCTPIALLVAGV----- 375
Db 3626 TDTSNNTQSTFSFALVTAGADATYLDATDILVTAAQAATTNINFTAADVVTVDTSN 3685
Qy 376 -----TGLISGLEASKQAMFESVANRLOKLEWEKQONGQNYFKGYDSRYAAYLAN 429
Db 3686 IQTNITITILAKVDKLSASSALDLTDSVTGKTVDV--SGVGDPTTITWAYDSSNVAF--- 3740
Qy 430 NLKFLSELNKELEAERVIAITQQRWNNNIGELAGITKLGRIKSGKAYADAFEGKKVEA 489
Db 3741 -----GSLTSTGGGELILSVT-----GTGTLATVTGL-----SNFNOLSA 3775
Qy 490 GSNITLDA-----KTGIIDISNSNGKKTQALHFTSPLLTAGTESRRLTNGKYSYI 540
Db 3776 ASDLTIDSAQVAGQTLMDSGAGNVTVSLASATTSYSSLSLTSTGAGTGLQIGSTGTYT-- 3833
Qy 541 NKLKFGVRKMQVTDGSEASSKLDFSKVRIORVAETGTEIGLI VNAKAGNDDIFVGGQKM 600
Db 3834 -----SVTGWEIFDSITS--DVA-----ITVAASA VTKTVAFSGTGAV 3870
Qy 601 NIDGGD-----HD-----RVFYSKGGFGFNITVDGTSATBAGS 634
Db 3871 TIDGASSNATFLNVTHDSGTGLILAVSSSTGSVAFSSGAANFVKVSDGLTYTGITS 3930
Qy 635 YTVNRKVARGDIYHEVVVKROETKVKRTEITQYRD---YELRKVGYGQYQSTDNLKSVEE 690
Db 3931 -----ANG-----QIIVNGSTSVAGISGTAYTDFIWNVDTSGLTYASTQGLL----- 3975
Qy 691 VIGSQFNDVFKSGKFNDIFHSGEGDLDLDGAGDDRLFGGKGNDRLSGDEGDDLLDGGSG 750
Db 3976 -----GGAGDDRL-----IDTSSGVLLAGEAG 3998
Qy 751 DDVLNGAGNDVYIFRKGD--GNDTLYDGTGNDKLAADANISIMIERKTEGIIVKEND 808
Db 3999 DDTLTGGAGNDIF--YHDGSSHGLTITTFASGDLQLQIASGNAGOWTLDSTGQALTKATNS 4057
Qy 809 HSG 811
Db 4058 SLG 4060

RESULT 9

US-10-238-075-749
; Sequence 749, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
; APPLICANT: I.N.S.E.R.M.
; TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolat
; TITLE OF INVENTION: E.coli, and biological uses of these polynucleotides and of thei
; FILE REFERENCE: BLANDINE
; CURRENT APPLICATION NUMBER: US/10/238,075
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 0003145
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 1576
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 749
; LENGTH: 1778
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-238-075-749

Query Match 5.0%; Score 232; DB 12; Length 1778;

Best Local Similarity 20.7%; Pred. No. 2.1e-07;
Matches 214; Conservative 134; Mismatches 417; Indels 268; Gaps 45;

Qy 6 VIKSNIQAGLNSTKSGLNKLYLAIPKDYDP-----OKGGTLDNDFIK---AADELG 52
Db 36 LISALVAGLLSFGSADNVTGQPTYDGSGAGDGVVAIGKAKANTFMTNSGASTALG 95
Qy 53 IARLAEPNNHTETAKKSVDTVNQFLSTQTGTGIAISATKLEKFLQKHSTNKLAKGLDLSVEN 112
Db 96 YDAIAEGEYSSAIGSKTLATGGASMAF---GVSAKAM-----GDRSVALGASSVAN 143

QY 113 IDRLGRASNVLSSTLSFLGTALAGIELDSILKKGDAAPDALAKASIDLINELIIGLSQS 172
DB 144 GDRSM--AFRYAKTNGFTSLAIG-----DSSLADGK-----TIALGNATAKA 184
QY 173 TQTIEAFSSQALKLGSTISQAKGFSNIGKLNQNFSTKTLNGLBIIITGLLSGISAGFALA 232
DB 185 YEIMSIALGDNVANASKEYAMALGASSKAGGADSLAFGRKSTANS--TGSALA-IGADSSSS 241
QY 233 DKNA-STGKKVAGFELSNQVIGNTVKAISYVLAQRYAAGLSTTGAAVALITSSIMLAI 291
DB 242 NDNAIAGNKTQA--LGVNSMALGNASQASGESSIALGNTSEASQNAI--ALGOGSIAASKV 299
QY 292 SPLAF--NNAADKFNHALDE-----PAKQFRKFCYDGDHLLAEVQRCVGTIEA 339
DB 300 NSIALGNSLSSGENALALGEGSAGGSNSLAFQSQSRANGNDVAI-----GVGAAAA 353
QY 340 SLTTISTALGAVSAGVSAAGV--SAVGTPIALLVAGV-----TGLISGILEASKQAMPES 393
DB 354 TDNSVAIGAGSTTDSATNTSVSGNSATKRKIVNMAAGAIISNTSTDAING---SOLYTTISDS 410
QY 394 VANPLOGKILEWEKONGQNYFDVG-----YDSRYAAYLANNLKFLSELNKELEABERVI 447
DB 411 VAKEL-----GGGATVGSDDGTAVSYALRSCTY----- 439
QY 448 AITQORWNNIGE--LAGITKLGRIKSGKAYADAFEDGKKVEAGSNITLDAK-----T 499
DB 440 -----NNVGDALSGID--NNTLQWNT--AGAFSAHGANATKNTNVAKGTVSATST 488
QY 500 GIIDISNSNGKKTQALHFTSPLTA--GTBSRRLTN-----GKYSYINKLKFRGVK 549
DB 489 DVNGSOLYDLQDALLWNGTAFSAHGTSEATSKITNTAGNTAGSTDAVNGSQL--KTT 547
QY 550 NWQVT-----DGEASKLDFSKVIQRYAETEGTDEIGLIVNAKAG 589
DB 548 NDNVTTNTNTIATNTNTITNLTDVAVGLGDDSLWKAAGAFSAHGTSEATSKITNTVNTAG 607
QY 590 N-----DDIFVGGKKNID-----GDDGHDRVFSYKGGFG 620
DB 608 NLTAGSTDAVNGSQLKTTNDNVTTNTNTIATNTNTITNLTDVAVGLGDDSLWKNKTAGAF 667
QY 621 NITVDGTSATE-----AGSYTVNRKVARGDYHEVVKRQETKVKRKTETIQRYDYELRK 674
DB 668 S-AAHGTDATSKITNTVAGNTAGSTDA---VNGSQLKTTNDNVTTNTNTIA----- 715
QY 675 VGYGYQSTDNLSKVEEVIGSFQFNDVFKSGFNDIFHSGEGDD-----LLDGGAGDRLFGG 730
DB 716 -----TNTNTITNLTDVAVGLGDDSLWKNKTAGAFSAHGTSDATSKITNTVAKAGD--LTAG 768
QY 731 KGNDRLSGDEGDDLLDGGSGD-----DVLNGAGNDVYIFRKGNDTLYDGTG--- 779
DB 769 -STDAVNGSQLKTTNDNVSTNTNTITNLTDVAV--GLGDDSLWKNKTAGAFSAHGTDATS 826
QY 780 ---NDKLAFADANISDIEMIERTKEGIIYKXRNDSHSGSINIPIRWYITSLNQYQSNKTDHKI 836
DB 827 KITNVKAGDLTAGSTDAV-----NGSQLKTTNDNVSTN-----TTNTITNLTDVSGDLKD 875
QY 837 EQLIGKDGSYITSDIKILQDKKQGVITTSQELKKLADENKSKOKLSASDIASSLNKLVG 896
DB 876 DSSLWNAKAAGAFS-----AAHGTEATS-----KITNLLAGKISSNSTDAIN 916
QY 897 SMALFGTANSVSS 909
DB 917 GSQLYGVADSFPS 929

RESULT 10

US-10-369-493-20175
; Sequence 20175, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374

; TYPE: PRT
; LENGTH: 499

; ORGANISM: No. US20030233675altoc punctiforme
US-10-369-493-20175

Query Match 5.0%; Score 230.5; DB 12; Length 499;
Best Local Similarity 23.0%; Pred. No. 3.9e-08;

Matches 122; Conservative 59; Mismatches 169; Indels 181; Gaps 24;

QY 409 NG--GQN--YFDKGYDSRYAAYLANNLKFLSELNKELEABERVIATQORWNNHIGELAGI 464
DB 58 NGEENNSLYGEAGNDSFYAANSNGN-----NLLDGGAGN 92

QY 465 TKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTG--IIDISNSNGKKTQALHFTSPLL 522
DB 93 DELVTVSSTG-----NNTLKGSGNDTLDISFSYGN-----IL 126

QY 523 TAGTESRERLTKGYSY--INKLKFGRVQWQVTD-----GEASSKLDFS---KVI 568
DB 127 DGSSED--DRL--SARYAVGNNTLRGGTGNDYIAADLSAGNNVLDGGDNDTLDINYQWKGK 184

QY 569 QRYAETEGTDEI-----GLIVNAKAGNDIDFVGQOK-----MNIIDGGGHDRVYF 613
DB 185 NTVSGGNGADTFYAYGVQGANITLNGSDSDSFYISSPETVLYSLVTQTVDDGGTNDYLYI 244

QY 614 SKDGGFENITVDGTSATEAGSYTVNRKVARGDYHEVVKRQETKVKRKTETIQRYDYELR 673
DB 245 DYSNATASITSTNPTNTNQLITAD-----TNQVRYKNIERL 281

QY 674 KVGYGYQSTDNLSKVEEVIGSFQFNDVFKSGFNDIFHSGE--GDDLLDGGAGDRL----- 727
DB 282 -----EVKGTAYDDNIVGSGDDTLLNGGNGNDTLKGGAGNDILRINDN 325

QY 728 -----FGKGNDRLSGDB--GDDLLDGGSGDD-----VLNGGAGNDV 762
DB 326 SRNTNTAYGCTGNDFLYAEATGCTSLDGGDGNLDYLSIANSYSVFGSRNTLNGVGNDT 385

QY 763 YIFRKGNDTLYDGTGNDKLAFAAD---ANISD-----IMIERTEGIIIVKR 806
DB 386 LNAAKATYVGNFLNGGVGND--FYADYSSGTNILDGDRDYLDSAGASFFHDFNTFTFLKG 444

QY 807 NDHSGSINIPRWYITSLNQYQSN-----KTDHKEIQLIGKGS--YITSDQ 851
DB 445 GTGNDITLANSFSFTNLDGCTGNDRLTVENSYKNNLLNGGNDYLLKGGQ 495

RESULT 11

US-09-813-214A-9
; Sequence 9, Application US/09813214A
; Patent No. US20020177200A1

; GENERAL INFORMATION:
; APPLICANT: Tucker, Kenneth

; APPLICANT: Plosila, Laura
; TITLE OF INVENTION: MORAXELLA CATARRHALIS OUTER MEMBRANE PROTEIN-106 POLYPEPTIDE, GENI

; FILE REFERENCE: 7969-089-999
; CURRENT APPLICATION NUMBER: US/09/813,214A

; CURRENT FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 08/968,685

; PRIOR FILING DATE: 1997-11-12
; NUMBER OF SEQ ID NOS: 20

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 2122
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-813-214A-9

Query Match      4.9%; Score 226.5; DB 10; Length 2122;
Best Local Similarity 20.3%; Pred. No. 6.9e-07;
Matches 227; Conservative 130; Mismatches 362; Indels 397; Gaps 54;

Qy 14 GINSTKGLKXLYLAIPKDYDPQGGTLNDRIKKADELGIARLAEEPHHTAKK-----S 69
Db 639 GINAGGKELTNVQSAI-----NPAITNGGQDPM-----NRLSTANTEKSGS 679

Qy 70 VDTVNFQSLTQTGTIAISATKLEFLQKHSTNKLAKGLDSVENIDRKLKASNYLSTLSS 129
Db 680 AATIKDLYLSQVPLTFAG-----DTGPNVTKKLGEILKV-----714

Qy 130 FLGTALAGIELDSLIIKGGDAAPDALAKASIDLI-----NEIIGNLSQSTQTIIEAFSSOLA 184
Db 715 -----KGGKTTADDLTKNIGVADSTDSNLTVKLAKTSLDLDVAVNTKL 759

Qy 185 KLGSTISQAGFSNIGKLN-----LNFSTKNLGLIIT-----GL-----LSGI 225
Db 760 TASDKVTVDSG-----NNTAKLQNGDLTFSTKONTGATPATNSKITGVNGLKFTDNNGIALDGT 817

Qy 226 S-----AGFALADKNASTGK-----KVAAG-FELSNQVIGNVTKAISSYVLAQRVAAG 272
Db 818 TYITKQKVGFAKQDGLSKPYLDKDKLVGEVEIITNGINAGGKAIT-----G 867

Qy 273 LSTTGAVAAITSSIMLAISPLAFMNAADK-----FNHVALDEFKAPQF--- 316
Db 868 LSNLTLDATNATTG-----HVTQLGIVDSTDKTRAASIGDVLNAGFNLNKNGD--AKDFVST 922

Qy 317 -----RKFGVGDHLLAEYQGV-GT-----IEASLTITSTA 347
Db 923 YDVTDFINGNATTAKVTYDGGKASVAYDVNVDGTTIHTGTADGNKNQOIGVKTTLTKTDA 982

Qy 348 LG-----AVSAGVSAAGVSAVGPPIALLVAGVTGLISGILEASKQAMFESVANRLOGK 401
Db 993 KGDKAINFVNSGDDKALIN-----AKDIADNLNTLAGEIRNTKGTADTADTFQVVK 1035

Qy 402 ILEWKQNGQNYFDKGVDSRYAAVLANNLPLSELNKELEAERVIATQORWNNIGEL 461
Db 1036 VKENGDDNDADTITVGKDAK--TNQVNTLKLKGNGLDIOQNK-----1077

Qy 462 AGITKLGRIKSGKAYADAFEDGKVEAGSNITLDAKTGIIDINSNGKKTQALHFTSPL 521
Db 1078 DGTVTFGINTQSG-----LKAGNNTTL-----NNGLSTKN-----1108

Qy 522 LTAGTESRERLTNGKSYINKLFGKRVKNQVOTDGEASSKLDFFSKVIQVIAETEGTDEIG 581
Db 1109 -TAGNEQIQVGADG-----VKPAKVN-----GVVGAGIDGTTITR-----DEIG 1148

Qy 582 LIVNAKAGNDIFVQGGKNIDGGCHDRVYKDGFGNITVDGTSATEAGSYTV-----637
Db 1149 F-----ACTNG-----SLDKSPH-----LSKDG-----INAGGKKTINIQSIEAONS 1187

Qy 638 NRKVARGDIIYEVVKRQETKVGKRTETIQYRDYELR-----KVGYGYSQSTDNLKS 688
Db 1188 NDAVTGGKIY-DLKTLELNKISSSTAKTQNSLHEFSVADEQGNFTVSNPSSYDTSKTS 1246

Qy 689 BEV-----IGSQFN-----DVFKG-----SKFNDFIHSQEGDLDGGGDDRLFG- 729
Db 1247 DVITPAGENGITTKVNGKVGVRVIGIDTKGLTTPKLTGVNNGKG-IVIDSQNGQNTITGL 1305

Qy 730 -----GKGNDRLSGDEG-----DILLGGSGDDVINGG-----AGN-----DV 762
Db 1306 SNTLANVTNDKSGVRIT--EQGKI IKDEKTRAASIVDVLNAGFNLCQNGEAVDFVSTYDT 1364

Qy 763 YIFKRGDNDT-----LYDGTGNDKLAFAFADANISDIMIE--RTKEGI-----802
```

```
1365 VNF--ADGNATTAKVTYDDTSKTSKVYVDVNVDDTTIEVKDKKLGVKTTTLTSTGTGANK 1422
803 -----IVKRNHSGSINIPRWYITSNLQNYOSNKTDRHKIEQLIKDGSYITSDQ 851
1423 FALSNOATGDALVKASDIVAHLANT-----LSGDIQTAKGASQANSAGYVDADGNKVIYDS 1478
852 ID-KILOQKDGTVITTSOELKK-----LADENKSOKLS 883
1479 TDNKYQAKNDGTVDKTKVEAKDKLVQAQTPDGTGLAQMNVKSVINKEQVNDANKOGIN 1538
884 ASDI-----ASSLNKLVGSMALFGTANSVSSNAL 912
1539 EDNAFVKGLEKAAADNKTNAAVTVGDLNAVAQTEL 1574

RESULT 12
US-09-932-183A-2
; Sequence 2, Application US/09932183A
; Patent No. US20020127641A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC394C1-US
; CURRENT APPLICATION NUMBER: US/09/932.183A
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/308,375
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/US98/18828
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: EP9719636.4
; PRIOR FILING DATE: 1997-09-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2285
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-932-183A-2

Query Match      4.7%; Score 219; DB 10; Length 2285;
Best Local Similarity 20.6%; Pred. No. 2.7e-06;
Matches 225; Conservative 153; Mismatches 375; Indels 340; Gaps 55;

Qy 4 INVTKSNIQAGLNSTKSGLEKNLYLAIPKDYDPQGGTLNDRIKKADELGIARLAEEPHHT 63
Db 366 IDTLMTNIRRVMSBPDYKYNEL-LQESIDLGTLNKNITDILQMTGDFG--RMGDFSESL 422

Qy 64 ETAKSVDTVNFQSLT--QTGTAISATKLEFLQKHSTNKLAKGLDSVEN-----ID- 114
Db 423 STLTQTAQVQLQVNSDLTPDDTVNTLTAAMLFNIAANDSISIAADKLAENVNNAVATTLDL 482

Qy 115 ----RKLKASNVSTLSSFLGTALAGIELDSLIIKGGDAAPDALAKASIDLINEIGNLS 170
Db 483 ANSRKAG-----STASTF-----GVELNDLI-----GYTTAIASTTRESGN-IVGN-- 523

Qy 171 QSTOTIEAFSSQLAKGSTISQAGFSNIGKLNQNFSTKNLG-LEIITGLLSISAGF 229
Db 524 -SLKTI-----FARIGNNOSSIKALEQIGISV-----KTAGGEAKSASOLISEVAGKW 570

Qy 230 -ALAD-----KNASTGKVAAGFELS--NOVIGNVTKAISSYVLAQORVA-AGLSTTGAV--- 279
Db 571 DTLSDAQKQNTSIG--VAGIYQLSRFNAMNN-----FSIAQNAAKTAANSTGSAWSE 621

Qy 280 -----AALITSSIMLAISPLAFMNAADKFNHVALDEFKAPQPRK 318
Db 622 QOKVADSLQARVNKLQNNFTFAAASDAFISDGLIEFTQAA-----GSLINASTGVIKS- 676

Qy 319 FGVDGDIHLAE-----YQRGVGTIEASLTITISTALG-----AVSAGVSAAGVSAV- 364
Db 677 VGFPLPPLAAVSTATLLSKNTRTLASSLILGTITAMGOETLATAGLEAGMTAAVASRVL 736

Qy 365 ----GTPIALLVAG-----VTGLISGILEA-----SKQAMFESVANRLOG-- 400
```

Db 737 KTAGLLVSTLVGFAALGMALESLSIPAEAKKAKDPEQSQQTNVRAITTNKQSTD 796
Qy 401 KILEWEKQ-----NGGQNYFD-----KGVSRYAYAVLANLKLFLS 435
Db 797 KLIQOYKELQVKESRLTSDEQEYLVQVTOQLAQTFPALVKGVDQSOGNAILKTN-----K 852
Qy 436 ELNKELEAERVIAITQQRWNNIGELAGITKLGRIKSGKAYADAFEDGKKVEAGSNITL 495
Db 853 ELEKAIE-----NTKEYLALKQETRDRAKTFEDASKEIK-----889
Qy 496 DAKTGIIIDISNGKKTQALHFTSPLLTAGTESRLTNGKY--SYINKLKFGVRVQWQV 553
Db 890 -----SKDELKQYKQIADYNDK---GRPK-WDL 913
Qy 554 TDGASSKLDPSKVIQVRAETEGTDEGLIVAKAGNDDIPVGGOKNIDGGDGHDRVY 613
Db 914 IADDDYKVAADKAKQSMKQAQSDIESG---NAKVKDSVLISANAYSSID-----960
Qy 614 SKDGGFNGITVDGTSATEAGSYTVNRKVARGDIYHEVVVKROETKVGRKTETIQYRDVELR 673
Db 961 -----ISNTLKTSI--SDVVKLNKLDLDPBELEKFSLSGLKQEKWQ-----K 1003
Qy 674 KVGYGQST-DN-----LKSVEEVIGSQFN--DVFK-----GSKPNDIFHSQEG 714
Db 1004 ALDSGDEKAFDPAKQDLQSLLETYSKSDSISDVFKMSFDKAKQNKIDGDKSLSVKSEVG 1063
Qy 715 D--DLLOGGAGDRLFGKGNDRLSGDEGDL-----LOGSGDDVLNGGAGNDV 762
Db 1064 DLGETLAEGNEAEDFGKLEALDANSVDIDKAAIKEMSDAMQFDSVQDVING-----DI 1119
Qy 763 YIPKRGDQ--NDTLTDTGNDKLA--PADANISDTMIERTKEGIIVKRNHSGSNIP 816
Db 1120 FNNTKQVAPLNDLL-----EKWAEKSIISANEANTLIQDKRE-LAQASIEGUVVKIN 1172
Qy 817 RWYITSNLQYQSNKTDHRIQLICKQDGYTSIQI---DKILOQKDKGVIT-----866
Db 1173 R-----DEVIKQKVLDAY--NDWVTYSNKLKMTVEVNAIKTLNADTLR 1215
Qy 867 SQELKULADENKSKLSA-----SDTASSLNKLUGSMALFGTANSVSSNALQ 913
Db 1216 IDSUKLKRKERKLDWSEALSDELVKSINNVAADAKELKLEKMLQPGGYNSQIEAMQ 1275
Qy 914 PITQPTGILAPS 926
Db 1276 SVKSALESYIISAS 1288

RESULT 13

US-10-192-584-7
Sequence 7, Application US/10192584
Publication No. US20030027987A1
GENERAL INFORMATION:
APPLICANT: TOKUNAGA, Eiji
SAKAGUCHI, Masashi
MATSUO, Kazuo
HAMADA, Fukusaburo
TOKIYOSHI, Sachio
TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS
PARAGALLINARUM AND PROCESS FOR PREPARING THE SAME
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 624 Ninth Street, N.W., Suite 300
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/192,584
FILING DATE: 11-Jul-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/077,098
FILING DATE: 19-May-1998
APPLICATION NUMBER: PCT/JP97/03222
FILING DATE: 12-SEP-1997
APPLICATION NUMBER: JP 27,148/1996
FILING DATE: 19-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: KORNBAU, Anne M.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TOKUNAGA=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2039 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-192-584-7
Query Match 4.6%; Score 216; DB 15; Length 2039;
Best Local Similarity 20.2%; Pred. No. 3.7e-06;
Matches 215; Conservative 126; Mismatches 424; Indels 298; Gaps 44;
Qy 11 IOAGLNSTKSLGNLYLAIPKDYDPQKGGTLPNDPIKAADBLGIALAEAP-----60
Db 337 IAGENSTAVGSLAFALA-----DKSTAMGLRSFV-AKDVGTAIGEESRTFAKDSVAI 390
Qy 61 -NHTETAKKSDVTNPFSLTQTG--IAISA-----TKLEKFLQKHSNKLAKGLD 108
Db 391 GNKTEASNAGSMAYGYKAKAVGAGAIAGAEGAFAEDSSQAGNLLNLRGAYATLKSD 450
Qy 109 SVENIDRKLGKASNVLSL-----SSFLGTALAGIELDSLKKGDAPDA 153
Db 451 KSDDI--KAGDAINVTFQFPDNMLTQGSHTYENTYLTTSAGDIK-KTLAAVGDGKNA 507
Qy 154 LAKASIDLINEIIGN--LSOSTOTIEAFSSQLAK-----LGS-----TISQAK 194
Db 508 IA-----IGNKTFASKANSVALGVALASQAQNALGVALGVALSPLAANTIVIGV 557
Qy 195 GFNSIGNKLNQNFSTNLGLEIITGLLSISAGFALADKNASTGKKVAAGFELSNQVIG 254
Db 558 GYATGNSFVGGSWVSTLSARTVVLGYSASIS-----SDSHDSLAMGV-----NAFIG 605
Qy 255 NVTKATSSYVLAORVAAGLSTTCGAAALITSSIMLAISPLAFMNAADKFNHA-----NALD 310
Db 606 NGSNS-----SLALGTGTIAKNAKSPSLAIGKDSRIDAKDITDNGVLTPQVYD 655
Qy 311 EFAKQPRKFGYDGDHLL-----AEYQRGVGTIEASLTTI-----STALGA 350
Db 656 ETTRAFATFDENKDYRQAMALGFNAKVRGKGMETGINSMAIGARSQATLQNSTALGV 715
Qy 351 -----VSAGVSAAGVAVGTPIALLVAG-----VTGLISGIL--EASKQA 389
Db 716 NAKTDYTWEOLEADPWVSKGAISIPTSKGIGVISVSGSKGERRIVNVASGLDITDANVA 775
Qy 390 MFESVANRLQKILEWEKQNGGQNYFD-----KGYDSRYAAYLANNL 431
Db 776 QLKTIERFQSEI--DLLQNGGGVQYLSVEKTNINGEAGRVASQIRKESGYKRYVKLKTQL 834
Qy 432 KFLSELNKELEAERVIAITQQRWNNIGELAGITKLGRIKSGKAYADAFEDGKKVEAGS 491
Db 835 LYL--DARKKNGEKFDOTSLDKISKAVQEL-----EAYSSELKTTASELNRVA---882
Qy 492 NITLDAKTIIDISNSNGKKTQALHFT-----SPLLTAGTESRLTNGKYSYI 540
Db 883 -MQLNAETTVDGFKFNQYKTIENATNADSEKNVGLSPQVIAQLKANNNNYLDGAKGQ 941


```
Qy 776 DGTGNDKL 783
Db 118 GGVGNLIL 125

RESULT 17
US-09-841-786-1
; Sequence 1, Application US/09841786
; Patent No. US20020054883A1
; GENERAL INFORMATION:
; APPLICANT: NAGARAJA, T. G.
; APPLICANT: STEWART, GEORGE C.
; APPLICANT: NARAYANAN, SANJEEV K.
; APPLICANT: CHENGAPPA, M. M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
; FILE REFERENCE: 30296
; CURRENT APPLICATION NUMBER: US/09/841,786
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 09/558,257
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3241
; TYPE: PRP
; ORGANISM: Fusobacterium necrophorum
US-09-841-786-1

Query Match 4.3%; Score 199; DB 9; Length 3241;
Best Local Similarity 21.1%; Pred. No. 0.00012;
Matches 194; Conservative 132; Mismatches 362; Indels 230; Gaps 44;

Qy 5 NVIKSNIQAGLNSTKSG-----LKNLYLAIPKDYDPQKGGTLNDFIKADELGIARLAEAP 60
Db 2382 NIKQTAANAGIGGAIGANVLNNPGTAVEDRKNSGKGT--EVLKTLDEV-----2430

Qy 61 NHTETAKSVDTNVQFLSLTGTGIAISATKLEK-----FLQKSTNKLAKGLDSV-ENDRK 116
Db 2431 -NKEODKVNATKKILO--SAGISTEDTSVKADRGDTQGEIGKAIVKTSIIIGKNVDIT 2487

Qy 117 LKASNVLS-----LSSFLGT--ALAGIELDSLIIKKGDAAPDALAKAS--IDLNEII 166
Db 2488 TEDKNITSTGGLGTAGLASAGTVAIVNIKNKNS-----GVTVENSFVAAEKVNRSDIT 2543

Qy 167 GNLSQSTOTIEAFSSQLAKGSLTISQAK-----GFSNI--GNKLQNLNF-----SKNLG 214
Db 2544 GNVA-----LTAYQGPVGAIGIGAAYAEIINSNGRSNISIKNSKLLGKNIDIVVKDKSEL 2598

Qy 215 LEITGILLSGISAGPALADK-----NASTGKKVAAGFELSNOVIGN 255
Db 2599 AE-AKGLTVGAAGAIIISKAKNMENSEVEIEKSIFFNEENRVTPSPKGIGREI-----N 2651

Qy 256 VTKAISVVLAORVAAGLSTTGAVA-ALITSSIMLAISPLAFMNAADKFN---HANALD 310
Db 2652 VKVEENRVTAESQQA---SVGAVAGAGIISEAKDAGS--SYLKVSTKSGRSIFHADNVN 2706

Qy 311 EFAKOFKFGYDGDHLLAEYQGVGTIEASLTTISTALGAVSAG-----VSAAAVGSVAG 365
Db 2707 MEATHKMKVTVASKAVTGSVLGGVGVTKAEATAAGKTWVEEENLFRTNRLNAISKVEG 2766

Qy 366 -----TPIALLVAGVTGLI--SGILEASKQMPESVANRLQKILEWEKQNGQNYFDK 417
Db 2767 LDEKVTAKSSVSGGGGIAGAGYNTSTAQNTESV-RLR-----KODYENNNDYTK 2818

Qy 418 GYDSRYAAYLANNKFLSELNKELEAERVIAITQORWNNIGEL--AGITKLGERIKSGK 475
Db 2819 -----KYISEVN-----ALALNDTKNEANTESLAVAGVHAQ-----TNK 2853

Qy 476 AYADAFEDGKKVAGSNIT-LDAKTGIIIDISNSGKKTQALHFTSPLLTAGTESERLTN 534
Db 2854 APTRSNKLSTTVNGGNSQLRAKALAKNENYGNVKGTT-----GGALVGAETAAYENYTK 2908

Qy 535 GKYSYINKLKFGRVKNQVTDGEASSKLDKFSKVIQORVAETEGTDEIGLIV-----584
Db 2909 ---STTCALVAG---NWEIGD-----KLEIARDNTIVRVNGDGTGKGLVKNGISVKNT 2957

Qy 585 ---NAKAGND--IFVQGGKMNIDGGDHRVFFYKDGFGFNITVDGTSATEAGSYTVNR 639
Db 2958 ISGETKSSIEDKARIIVGTGSVNVDALELDVLDLQKSGGYYGIGIGNV-----3005

Qy 640 KVARGDIYHEVVKRQETKVKRRTETIOYRDYELAKVGYGQSTDNLKSVEVIGSQNDV 699
Db 3006 ---DVNNVIKKVNEAKIGRH-----AIVETTKQOEYQA 3035

Qy 700 FKSGKFNDIFHSGEGDLDLGGAGDRLFGKGNDRLSGD-EGDDLLDGGSGDDVLANGGA 758
Db 3036 FTRAKVNIL-----GKGDAAAAAAISNVHISNEMDIKNLAKQYASSQLITQNS 3083

Qy 759 GNDVYIFRKGNDLYDGTGNDKCLAFADANIS--DIMIERTKEGIIIVKRNDSHSGSINIPR 817
Db 3084 KNNITLASSSESNNVNH-GVAEARGAGAKATVSVKQINRNTNNVDLAGIKTEGINVYA 3142

Qy 818 WYTSNLQNTQSNKTDHK 835
Db 3143 GYD-----KNYNISKTSK 3156

RESULT 18
US-10-175-275-4
; Sequence 4, Application US/10175275
; Publication No. US20030171254A1
; GENERAL INFORMATION:
; APPLICANT: SASAKI, Ken
; APPLICANT: HARKNESS, Robin E.
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF
; TITLE OF INVENTION: MORAXELLA
; FILE REFERENCE: 1038-1235 MIS
; CURRENT APPLICATION NUMBER: US/10/175,275
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 08/945,567
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 08/431,718
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: 08/478,370
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/621,944
; PRIOR FILING DATE: 1996-03-26
; PRIOR APPLICATION NUMBER: PCT/CA96/00264
; PRIOR FILING DATE: 1996-04-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1833
; TYPE: PRP
; ORGANISM: Moraxella catarrhalis
US-10-175-275-4

Query Match 4.2%; Score 197; DB 12; Length 1833;
Best Local Similarity 19.6%; Pred. No. 7.3e-05;
Matches 238; Conservative 157; Mismatches 407; Indels 412; Gaps 61;

Qy 2 SNINVIK-----SNTQAGLNSTKSLKNLYLAIPKDYDPQKGGTLNDFIKADELGIARLA 57
Db 181 NNIGVKEADNSGLKVKLAKTLNNLTENVTTTLNATTVTKVSGSSS--TTAELLSDSLTF 238

Qy 58 EEPN-HTETAKSVDTVN--QFLSLTOTGIAISATKLEKFLQKHSTNKLAKGLD-----108
Db 239 TOPNIGSQSTSKIVYGVNGVKFTNNAETAAIGTRITR-----DKIGFARGDVDEK 291

Qy 109 SVENIDRKLKGSANVLSLTSFLGTALAGIELDSLIIKKGDAAPDALAKASIDLINIEIGN 168
Db 292 QAPYLDKKQLKVGSVAITIDNGID---AGNKKISNLAKGSSANDAVT-----IEQ 338
```

```
QY 169 LSQSTQTIKAFSSQLAKLSTISQAKGFSNIGKLNQNLNFSKTNLGLLEIITGLLSGASG 228
Db 339 LKAAPTUNA-----GAGISVTPTEISVDAKSGNVTAPTNYIGVKTTELNSDGTSDK 390
QY 229 FALADKNASTGKVAAGFELS---NOVIGNVTKAISYVL-----AQRVAAGLSTT-- 276
Db 391 FSV--KSGTNNSLVTAHLASYLNEVNRVTSALQSFTVKEEDDDANAITVAKDTTKN 448
QY 277 -GAVAAL-----ITSSIMLAI 291
Db 449 AGAVSILKLGKNGLTVAATKDGTVTFGLSQDSGLTIGKSTLNNDGLTVKDTNEQIQVGA 508
QY 292 SPLAFMAADKFHANALDEFKQFR---KFGYDGDHLLAEYQGVGTIEASLTITSTA 347
Db 509 NGIKFTNV---NGSNPGTGIANTRITRDKIFGAGS-----DGAVDTKNPKYLDQDKLQ 558
QY 348 LGAV---SAGVSAAGVSAVGTPIALLVAGVTGLISGLEASKQAMPESVANRLQKILE 404
Db 559 VGNVKITNTGINAG--GKA-----ITGLSPTLPSIADQSSRNI--ELGNTIQDK--- 603
QY 405 WEKQNGG--QNYFDKGYDSRYAAYLANN---LKFLSE-----LNKELEAERVIAITQOR 453
Db 604 -DKSNAASINDILNTGFN-----LKNNNPPIFVSTYDIVDFANGNATTATVTHDTANK 656
QY 454 -----WNNNIGE---LAG-----ITKLERIKSGKAYAD-----AFEDG----- 484
Db 657 TSKVYVDVNVDDTTIHLTGTDNDKKGKVTTKLNKTSANGNTATNFNVNSDEDAVNAK 716
QY 485 -----KKEVAGSN-----ITL 495
Db 717 DIAENLNTLAKEIHTKTGADTALQTFVKKVDENNADANAITVQKKNANNQVNTLT 776
QY 496 DAKTGIIIDINSNGKKTQALHFTSPLTAGTESRRLTNGKYSYINK-----L 543
Db 777 KGENGLNIKDKNGTVFGINTTSG-LKAG---KSTLNDGGLSKNPTGSEQIQVGADGV 832
QY 544 KFGVKVQWVTDGASSKLFDSKVIQVVAETEGTDEGLIVNAKAGNDIIFVGGKKNID 603
Db 833 KFAKVN---NGVVGAGIDGTTTRITR-----DEIG-----FTG-----TN 864
QY 604 GGDGHRVFSKGGFGNITVDGTSAT--EAGSVTVNR--KVARGDIYHEVVKQETKVG 659
Db 865 GSLDKSRPHLSKDG---INAGGKKITNIQSGEATAQNSHDAVGTGKIY-DLKTLENKIS 919
QY 660 KRTETIQVRVELR-----KVGYGYQSTDNLKSVEEV-----IGSQFN----- 697
Db 920 STAKTAQNSLHEFSVADEQGNFTVSNPYSSYDTSKTSVDITFAGENGITTKVNGVVRV 979
QY 698 --DVFPKG---SKFNDIPIHSGEGDLDLGGAGDDRLFG-----GKGNDRLSGDBGD 742
Db 980 GIDQTKGLTTPKLVGNNNGKG-IVIDSQNGQNTITGLSNLTANVTNDKGSVRRT--EQGN 1037
QY 743 DLLD-----GGSGDDVLNGG--AGN-----DVYIFRKGDGNDT-----LYDGTGN 780
Db 1038 IIKDEDKTRASIIVDVLSAGFNLOQNGEAVDFVSTYDTVNF--ADGNATKATVYDDTSK 1095
QY 781 DKLAFADANISDIMIE--RTKEGI-----IVKRNDSHSGSI 813
Db 1096 TSKVYVDVNVDDTTIEVKDKLGKVTTLTSTGTGANKFALSNOATGDALVKASDIIVAH 1155
QY 814 NIPRWYTSNLOQNSKTDHIIQLIGKGSYITSDQID-KILODKDGTVITTSQELKK 872
Db 1156 NT-----LSGDIQTAKGASQANNSAGYVDADGNKVIYDSTDNKYYQAKNDGTVDKTEVAK 1211
QY 873 -----LADSNKSQKLSASDI-----ASSLNKLVGSM 898
Db 1212 DKLVAQAQTPDGTPLAQMNVKVINKEQVNDANKQKGINEADNAFVKGLEKAASDNKTNA 1271
QY 899 ALFGTANSVSSNAL 912
Db 1272 VTVDGLNAVAQTPL 1285
```

```
RESULT 19
US-10-175-282-4
; Sequence 4, Application US/10175282
; Publication No. US20030170657A1
; GENERAL INFORMATION:
; APPLICANT: SASAKI, Ken
; APPLICANT: HARKNESS, Robin E.
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: MORAXELLA
; FILE REFERENCE: 1038-1234 MIS
; CURRENT APPLICATION NUMBER: US/10/175,282
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 08/945,567
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 08/431,718
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: 08/478,370
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/621,944
; PRIOR FILING DATE: 1996-03-26
; PRIOR APPLICATION NUMBER: PCT/CA96/00264
; PRIOR FILING DATE: 1996-04-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1833
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-10-175-282-4
```

```
Query Match 4.2%; Score 197; DB 12; Length 1833;
Best Local Similarity 19.6%; Pred. No. 7.3e-05;
Matches 238; Conservative 157; Mismatches 407; Indels 412; Gaps 61;

QY 2 SNINVIK-----SNIQAGLNSTKSLGNLYLAIPKDYDPQKGTLDNFIKADELGIARLA 57
Db 181 NNIGVKEADNSGLKVLAKTLNLTNTTNNATTTTVKVGSSSS--TTAELLSDSLTF 238
QY 58 EEPN-HTETAKKSVDTVN--OFLSLTQTGTIAISATKLEKPLKHQSTNKLAKGLD----- 108
Db 239 TQNTGSSQSTKTVYGVNGVKFTNNAETTAIGTTRITR-----DKIGFARDGVDDEK 291
QY 109 SVENIDRKLGKASNVLSLSSFLGTALAGIELDSLICKGDAAPALAKASIDLINEIGN 168
Db 292 QAPYLDKKQLKVGSAVITDNGID---AGNKKISNLAKGSSANDAVT-----IEQ 338
QY 169 LSQSTQTIKAFSSQLAKLSTISQAKGFSNIGKLNQNLNFSKTNLGLLEIITGLLSGASG 228
Db 339 LKAAPTUNA-----GAGISVTPTEISVDAKSGNVTAPTNYIGVKTTELNSDGTSDK 390
QY 229 FALADKNASTGKVAAGFELS---NOVIGNVTKAISYVL-----AQRVAAGLSTT-- 276
Db 391 FSV--KSGTNNSLVTAHLASYLNEVNRVTSALQSFTVKEEDDDANAITVAKDTTKN 448
QY 277 -GAVAAL-----ITSSIMLAI 291
Db 449 AGAVSILKLGKNGLTVAATKDGTVTFGLSQDSGLTIGKSTLNNDGLTVKDTNEQIQVGA 508
QY 292 SPLAFMAADKFHANALDEFKQFR---KFGYDGDHLLAEYQGVGTIEASLTITSTA 347
Db 509 NGIKFTNV---NGSNPGTGIANTRITRDKIFGAGS-----DGAVDTKNPKYLDQDKLQ 558
QY 348 LGAV---SAGVSAAGVSAVGTPIALLVAGVTGLISGLEASKQAMPESVANRLQKILE 404
Db 559 VGNVKITNTGINAG--GKA-----ITGLSPTLPSIADQSSRNI--ELGNTIQDK--- 603
QY 405 WEKQNGG--QNYFDKGYDSRYAAYLANN---LKFLSE-----LNKELEAERVIAITQOR 453
Db 604 -DKSNAASINDILNTGFN-----LKNNNPPIFVSTYDIVDFANGNATTATVTHDTANK 656
```

454 QY -----WNNIGE-----LAG-----ITKLERIKSGKAYAD-----AFEDG----- 484
657 DB TSKVYDVNVDTTIHLTGTDNDKKGKVTTLAKNTSANGNTATNFVNSSDEDALVNAK 716
485 QY -----KVEAGSN-----ITL 495
717 DB DIAENLNTLAKEIHTTKGTADTALQTFVKKVDENNADANATVGGKANNQVNTLTL 776
496 QY DAKTGIIIDISNSNGKKTQALHFTSPLTAGTESRRLTNGKYSYINK-----L 543
777 DB KGENGLNIKTDKNGTVTFGINTTSG-LKAG---KSTLNDGGLSIKNPPTGSEQIOVGADGV 832
544 QY KFGVKNQVDTGEASSKLDPSKVIQVVAETEGTDEIGLIVAKAGNDDIVFGQKKNID 603
833 DB KPAKVN---NGVVGAGIDGTTTRITR-----DEIG-----FTG-----TN 864
604 QY GGDGHDVRFYSKDGFGNITVDGTSAT--EAGSYTVNR--KVARGDIVHEVVKRQETKVG 659
865 DB GSLDKSKPHLSKDG---INAGGKIITNIQSGETAQNASHDVTGKIIY-DLKTLENKIS 919
660 QY KRTETIQVRYELR-----KVGYGQSTDNLKSVEEV-----IGSQFN----- 697
920 DB STAKTAQNSLHEFSVADEQGNFTVSNPYSSYDTSKTSVDITFAGENGITTKVNGVVRV 979
698 QY -DVPFG--SKFNDFHSCEGDDLLDGGAGDRLFG-----CKGNDRSLSGDEGD 742
980 DB GIDQTKGLTTPKLTGVNNGKG-IVIDSQNGQNTITGLSNTLANVTNDKGSVRTT-EQN 1037
743 QY DLLD-----GSGDDVLNGG--AGN-----DVYIFRKGDNNDT-----LYDGTGN 780
1038 DB IKBEDKTRASIVDVLGAGNLOQNGEAVDFVSTYDTPNP--ADGNATTAKVIYDDTSK 1095
781 QY DKLAFADANISDIME--RTKEGI-----IVKRNHDSGI 813
1096 DB TSKVYDVNVDTTIEVKDKLGKVTTLTSTGTGANKFALSNOATGDALVKASDIVAHL 1155
814 QY NIPRWIYTSNLQYNSKTDHKIEQIIGKDSYITSDQID-KILODKKQGVITISQELKK 872
1156 DB ---LSDIQTKAGASOANNSAGYVDAGNKVIYDSTDNKYQAKNDGTVDTKEVAK 1211
873 QY -----LADENKSKLSASDI-----ASSLNKLVGSM 898
1212 DB DKLVAQAQTPDGTALQMNKVSINKEQYNDANKQGINEDNAFVKGLEKASDNKTKNAA 1271
899 QY ALFGTANSVSSNAL 912
1272 DB VTVDGLNAVAQTPL 1285

RESULT 20
US-10-175-275-3
; Sequence 3, Application US/10175275
; Publication No. US20030171254A1
; GENERAL INFORMATION:
; APPLICANT: SASAKI, Ken
; APPLICANT: HARKNESS, Robin E.
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF
; FILE REFERENCE: 1038-1235 MIS
; CURRENT APPLICATION NUMBER: US/10/175,275
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 08/945,567
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 08/431,718
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: 08/478,370
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/621,944
; PRIOR FILING DATE: 1996-03-26

; PRIOR APPLICATION NUMBER: PCT/CA96/00264
; PRIOR FILING DATE: 1996-04-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 3
; LENGTH: 1992
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-10-175-275-3

Query Match 4.2%, Score 197, DB 12, Length 1992;
Best Local Similarity 19.6%, Pred. No. 8.3e-05;
Matches 238; Conservative 157; Mismatches 407; Indels 412; Gaps 61;

QY 2 SNINVIK-----SNIOAGLNSTKSLKVLXLAIPKDYDPQKGGTLNDFIKKADELGIARLA 57
DB 340 NNIQVKEANSGLKVLAKTLNNLTENVNTTNNATTTVKVGSSS--TTAELLSDLTF 397
QY 58 EEPN-HTETAKSKVDTVN--QFLSLTOTGIAISATKLEKFLQKHSTNKLAKGLD----- 108
DB 398 TQPTNGSQSTKTVYGVNGVFTNNAETTAIGTTRITR-----DKIGFARDGVDEK 450
QY 109 SVENIDKLGKASNVLSLSTSLGTLALAGIELSLIKKGAAPDALAKASIDLINELIGN 168
DB 451 QAPYLDKKQLKVGSAITIDNGID---AGNKKISNLAKSSANDAVT-----IEQ 497
QY 169 LSOSTQTIFAFSSQAKLGSTISOAKGFSIGNKLNINFSKTNLGLLEITGLLSGISAG 228
DB 498 LKAAPPTLNA-----GAGISVTPEISVDKSGNVTAFTYINIGVKTTELNSDGTSDK 549
QY 229 FALADKNASTGKVAAGFELS---NOVIGNVTKAISSVYL-----AQRVAAGLSTT-- 276
DB 550 FSV--KSGTNSLVTAHSLASYLNEVNRATDSALQSFTVKEEDDDANAITVAKDTIKN 607
QY 277 -GAVAAL-----ITSSIMLAI 291
DB 608 AGAVSILKLGKNGLTVAKKDGTVTFLGSDSGLTIGKSTLNNDGLTVKDTNEQIOVGA 667
QY 292 SPLAFMAADKFHNANALDEFKQFR---KFGYDGDHLLAEYORGVGTTIEASTITISTA 347
DB 668 NGIKFTNV---NGSNPGTGIANTARITRDKIGFAGS-----DGAVDINKPYLDQDKLQ 717
QY 348 LGAV---SAGVSAAVGSAVGTPTIALIVAGVTGLISGLEASKQAMFESVANRLQGGKILE 404
DB 718 VGNVKIITNTGINAG--CKA-----ITGLSPTLPSIADQSSRNI--ELGNTIQQK--- 762
QY 405 WEKONGG--QNYFDKYDSRYAAYLANN---LKFLSE-----LNKLEAEARVTAIQOR 453
DB 763 -DKSNAASINDILNTGFN-----LKNNNNPIDFVSTYDIDVFANGNATTATVTHDTANK 815
QY 454 -----WNNIGE-----LAG-----ITKLERIKSGKAYAD-----AFEDG----- 484
DB 816 TSKVYDVNVDTTIHLTGTDNDKKGKVTTLAKNTSANGNTATNFVNSSDEDALVNAK 875
QY 485 -----KVEAGSN-----ITL 495
DB 876 DIAENLNTLAKEIHTTKGTADTALQTFVKKVDENNADANATVGGKANNQVNTLTL 935
QY 496 DAKTGIIIDISNSNGKKTQALHFTSPLTAGTESRRLTNGKYSYINK-----L 543
DB 936 KGENGLNIKTDKNGTVTFGINTTSG-LKAG---KSTLNDGGLSIKNPPTGSEQIOVGADGV 991
QY 544 KFGVKNQVDTGEASSKLDPSKVIQVVAETEGTDEIGLIVAKAGNDDIVFGQKKNID 603
DB 992 KPAKVN---NGVVGAGIDGTTTRITR-----DEIG-----FTG-----TN 1023
QY 604 GGDGHDVRFYSKDGFGNITVDGTSAT--EAGSYTVNR--KVARGDIVHEVVKRQETKVG 659
DB 1024 GSLDKSKPHLSKDG---INAGGKIITNIQSGETAQNASHDVTGKIIY-DLKTLENKIS 1078
QY 660 KRTETIQVRYELR-----KVGYGQSTDNLKSVEEV-----IGSQFN----- 697
DB 1079 STAKTAQNSLHEFSVADEQGNFTVSNPYSSYDTSKTSVDITFAGENGITTKVNGVVRV 1138

Qy 698 --DVPKG---SKENDIFHSCEGDDLLDGGAGDDRLFG-----GKNDRLSGDEG 742
Db 1139 GIDOTKGLTTPKLVGNNGKG-IVIDSQNGQNTITGLSNTLANVTNDKGSVRTT-EQGN 1196
Qy 743 DLLD-----GSGDDVLNGG---AGN-----DVIYFRKGDGNDT---LYDGTGN 780
Db 1197 IIKDEDKTRAASIVDVLGAGFNQNGEAVFVSTYDTVNF--ADGNATTAKVYDDTSK 1254
Qy 781 DKLAFADANISDIMIE--RTKEGI-----IVKRNDSHSGSI 813
Db 1255 TSKVYDVNVDDTTIEVKDKKGLGVKTTTLTSTGTGANKFALSNOATGDALVKASDI VAHL 1314
Qy 814 NIPRWYITSLNQNYSNKTQHKIEQLIGKDGSVITSDI-D-KILODKKDGVTITSOELKK 872
Db 1315 NT-----LSGDIOTAKGASQANSAGYVDAGNKVIYDSTDNKYQAKNGDGVDTKREVAK 1370
Qy 873 -----LADENKSQKLSASDI-----ASSLNKLVGSM 898
Db 1371 DKLVAQAOTPDGTLAQNPKVSKVINKQVNDANKKQGINEDNAFVKGLEKASDNKTNA 1430
Qy 899 ALFGTANSVSSNAL 912
Db 1431 VTVDGLNAVAQTPL 1444

RESULT 21
US-10-175-282-3
; Sequence 3, Application US/10175282
; Publication No. US20030170657A1
; GENERAL INFORMATION:
; APPLICANT: SASAKI Ken
; APPLICANT: HARKNESS, Robin E.
; APPLICANT: LOOMORE, Sheena M.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF
; TITLE OF INVENTION: MORAXELLA
; FILE REFERENCE: 1038-1234 MIS
; CURRENT APPLICATION NUMBER: US/10/175,282
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 08/945,567
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 08/431,718
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: 08/478,370
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/621,944
; PRIOR FILING DATE: 1996-03-26
; PRIOR APPLICATION NUMBER: PCT/CA96/00264
; PRIOR FILING DATE: 1996-04-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1992
; TYPE: PRF
; ORGANISM: Moraxella catarrhalis
US-10-175-282-3

Query Match 4.2%; Score 197; DB 12; Length 1992;
Best Local Similarity 19.6%; Pred. No. 8.3e-05;
Matches 238; Conservative 157; Mismatches 407; Indels 412; Gaps 61;

Qy 2 SNINVIK-----SNTQAGLNSTKSGLNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLA 57
Db 340 NNIGVWEADNSGLKVLAKTLNNLTNTVTTLNATTATTKVGSSES--TTAEILLSLTF 397
Qy 58 EEPN-HTETAKSVDTYN--QFLSLTQGIASATKLEKFLQKSTNKLKGLD-----108
Db 398 TQPTNGSQSTSKTVGVNGVRFNNAETTAAGTTRITR-----DKIGFARDGDVDEK 450
Qy 109 SVENIDRKLKASNVLSLTSSFLGTALAGIELDSLKKGDAAPDALAKASIDLINELIGN 168

Db 451 QAPYDKKQKLVGSAVITIDNGID---AGNKKISNLAKSSANDAVT-----IEQ 497
Qy 169 LSOSTOTIEAFPSQOLAKLGSSTISOAKGFSNIGNKLNLFNSKTNLGLIITIGLLSGISAG 228
Db 498 LKAAPFTLNA-----GAGISVTPTEISVDKSGNVTAFTYINIGVKTTELMSDGTSDK 549
Qy 229 FALADKNAASTGKKAAGFELS---NOVIGNVTYKAISYVL-----AORVAAGLSTT-- 276
Db 550 FSV--KSGGTNNSLVTAHSLASVYNEVRTADSALQSFTVKEEDDDDDANAITVAKDITKN 607
Qy 277 -GAVAL-----ITSSIMLAI 291
Db 608 AGAVSILKLGKGNGLVATVKDGTFTFGLSQDSGLTIGKSTLNNDDGLTVKDTNEQIQVGA 667
Qy 292 SPLAFNNAADKFNHANALDEFKQFR---KFCYDGDHLLAEYQRCGVGTIEASLITISTA 347
Db 668 NGKFTNV-----NGSNPGTGIANTRITRDKIFGAGS-----DGAVDTNKPVLDDQKLQ 717
Qy 348 LGAV---SAGVSAAAVGSVAGTPIALLVAGVTGLISGILEASKQAMFESVANRLQOKILE 404
Db 718 VGNVKITNTGINAG--GKA-----ITGLSPTLPSIADSSRNI--ELGNTIQDK--- 762
Qy 405 WEKONGG--QNYFDKGYDSRYAAYLANN---LKFLSE-----LNKELEAERVIAITQOR 453
Db 763 -DKSNAASINDILNTGFN-----LKNNNPIDEVSTYDIVDFANGNATTATVTHDTANK 815
Qy 454 -----WNNIGE---LAG---ITKGERIKSGKAYAD-----AFEDG----- 484
Db 816 TSKVYDVNVDDTTIHLTGDDNKKLGVTTKLNTKSANGNTATNFNVNSDEDALVNAK 875
Qy 485 -----KKVBAGSN-----ITL 495
Db 876 DIAENLTLAKEIHTTKGTADTALQFTVVKVDENNADANAITVQCKNANNQVNTLT 935
Qy 496 DAKTGIIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINK-----L 543
Db 936 KGENGLNIKTDKNGVTFTGINTTSG-LKAG---KSTLNDGGLSIKNTPTGSEQIQVGADGV 991
Qy 544 KFCRVKMWQVTDGEASSKLDKFSKVIQVATEGTDEIGLIVNAKAGNDDIFVCGKKNID 603
Db 992 KFAKVN---NGVVGAGIDGTTRITR-----DEIG-----FTG-----TN 1023
Qy 604 GGDGHRDVFYSKDGFGNITVDGTSAT--EAGSYTVNR--KVARGDIYHEVWRKQETKVG 659
Db 1024 GSLDKSKPHLSKDG-----INAGKKITNIOGELIAQNSHDAVTGGKIY-DLKTLENNKIS 1078
Qy 660 KRTETIYRDYELR-----KVGYGYSTNLKSVREV-----IGSQFN----- 697
Db 1079 STAKTAQNSLHEFSVADEQGNFTVSNPYSSYDTSKTSDVITFAGENGITTKVNGVVRV 1138
Qy 698 --DVPKG---SKENDIFHSCEGDDLLDGGAGDDRLFG-----GKNDRLSGDEG 742
Db 1139 GIDOTKGLTTPKLVGNNGKG-IVIDSQNGQNTITGLSNTLANVTNDKGSVRTT-EQGN 1196
Qy 743 DLLD-----GSGDDVLNGG---AGN-----DVIYFRKGDGNDT---LYDGTGN 780
Db 1197 IIKDEDKTRAASIVDVLGAGFNQNGEAVFVSTYDTVNF--ADGNATTAKVYDDTSK 1254
Qy 781 DKLAFADANISDIMIE--RTKEGI-----IVKRNDSHSGSI 813
Db 1255 TSKVYDVNVDDTTIEVKDKKGLGVKTTTLTSTGTGANKFALSNOATGDALVKASDI VAHL 1314
Qy 814 NIPRWYITSLNQNYSNKTQHKIEQLIGKDGSVITSDI-D-KILODKKDGVTITSOELKK 872
Db 1315 NT-----LSGDIOTAKGASQANSAGYVDAGNKVIYDSTDNKYQAKNGDGVDTKREVAK 1370
Qy 873 -----LADENKSQKLSASDI-----ASSLNKLVGSM 898
Db 1371 DKLVAQAOTPDGTLAQNPKVSKVINKQVNDANKKQGINEDNAFVKGLEKASDNKTNA 1430
Qy 899 ALFGTANSVSSNAL 912
Db 1431 VTVDGLNAVAQTPL 1444

RESULT 22

US-09-815-242-5639
; Sequence 5639, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; FILE FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5639
; LENGTH: 2086
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-5639

Query Match 4.2%; Score 196; DB 9; Length 2086;

Best Local Similarity 18.6%; Pred No. 0.0001;
Matches 221; Conservative 165; Mismatches 431; Indels 372; Gaps 54;

QY 1 MSNINVIKSNQAGLNSTKSGLNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
DB LEDIQVQTNGQA-LNNAMKGLRD---SIANE-----TTVKASQNYTDA----SP 701
QY 61 NHTETAKSVDT-----VNOFLSLTGTGIAISATKLEFLQKHSHTNKLAKGLDSVENIDRK 116
DB NQSTYNGAVSNAGKIINGNTNPTNMTDSITQATTQ-----VNNAKNGLNGAENLRNA 754
QY 117 LGKASNVLSLSTGFLGTAGLDELSLIKGDAAPDALAKASIDLINEIIGNLSQS----- 172
DB QNTAKQNLNTLSHLTNNOKSALS-SQIDRAGHVSEVTAAKAAATELNTOMGNLEQAIDHQ 813
QY 173 ---TQTIEAFSSQLAKGSTISQAKGFSNIGNKLQNLNFSKTNL--GLEIITGLLSGISA 227
DB 814 NTVKGQVNTDADKAKRDAYTNAVSRATEILNKTQANTSKQDVEAAIQNVTSKALN- 872
QY 228 GFALADKNAKSTGKKVAAGFELSNOVIGNVT-----KAISYVLAQVAVAG---LSTTG 277
DB 873 ----GDQVNTNAKNTA-----KHALNLT SINNAQKRDLTTKIDQATTVAGVAVSNTG 922
QY 278 AVAALITSLIMLAISPLAFMAADKFHNALDEFAPKFRKFGYDGDHLLAEYQKGVGT- 336
DB 923 TQLNTAMANLQINDKANTLASYNHDADSDKTTA--YTQAVTNAENIL---NKGSGN 977
QY 337 -----IEASLTITISTALGAVSAGVSAAGVAGTPPIALLVAGVTGLISGILEASKQAM- 390
DB 978 LDKAAVENALSQVNTAKGALNGHNHLEQAKSNANTTINGLQHLTTAOKDKLQVQQAQN 1037

QY 391 -----PESVANRLOG-----KILEWEKONGGQNYFD-----KGYDSRYAAYL 427
DB 1038 VAGVDTVKSSANTLNGAMGTILRNSIQDNTATNNGQNYLDATESKNTYNNNAVDANGVIN 1097
QY 428 A-----NNKFLSELNKE----- 440
DB 1098 ATSNPNMDANAINQIAQTQVTSKNALDGTNLTQAKQTATNAIDGATNLNKAQKDALKAQ 1157
QY 441 -LEAERVIAITO-QRWNINIGELAGITKLG-----ERIKSGKAYADAFEDGKKV----- 487
DB 1158 VTSQRVANVTSTQOTANELNTAMQLOHGIDDENATKQTKYRDA-EQSKTKAYDQAVA 1216
QY 488 -----EAGSN-----ITLDAKTIIDISNSNGKKTQAL-----HFTSPL 521
DB 1217 AAKAILNKQTSNDDKAADRALQOQVTSKDALNGDAKLAEEAKAAKQNLGTLNHIINAQ 1276
QY 522 LTA-----GTESRERLTNGKYSYINKL-----KFGRVKQMVQVTDGSSAKLD 563
DB 1277 RTALEGQINQATTVGQVNTVKTNANTLDGAMNSLQGSINDKDATLRNQNYLDADSKRNA 1336
QY 564 FSKVIQVATEGTDEIGLIVNAKAGNDDIFVGQKQWIDGGDGHDRVFSKD--GGFGN 621
DB 1337 YT---QAVTAEG-----ILNKQTGGN-----TSKADVD--NALNTVTRAKAALNGAEN 1380
QY 622 ITVDGTSATEAGSYTVN---RKVARGDIYVHVVRQ-----ETKVGKR 661
DB 1381 LRNTKTSATN-----TINGLNLTLQKONLKHQVEQAQNVAGVGVKDKGNTLNTAMGAL 1436
QY 662 TETIQ-----YRD-YELRKVGY----- 677
DB 1437 RTSIQNDNTTKTSQNYLDASDSNKQNYNTAVNNANGVINVTNPNMDANALNGMANQVNT 1496
QY 678 -----GYQ-----STDNLKSVEEIVGSQFN-----DVEKSGKEN 706
DB 1497 TKAALNGAQNLAQAKTAKNNLNLTSINNAQKDALTRSIDGATTVAGVNOETAKATELN 1556
QY 707 DIPHSGBGDDLLDGGAGDRLFGGKGNDRLSGDBGDDLLDGGSG-----DDVLNGGAGND 761
DB 1557 NAMHS-----LQNGI-----NDETQTKQTKYLDAGPSKSKAYDAVNAKA-- 1598
QY 762 VYIFRKGDNLTLYDGTGNDKLAFAADANI SDIMERTKEGIIVKRNDSHSGSINIPRWYIT 821
DB 1599 --ILTKASQNV-----DKAAVEQA-LQNVNSTKTALNGDAKLNEAKAAKQTLGLTL 1648
QY 822 SNLQYQSNKTDHKIEQLIGKQGSYITS-----DQIDKILQKQDQVITSQELK 871
DB 1649 -HINNAQRTALDNEITQATNVEGVNTVTKAKAQQLDGAQOLETSIRDK--DTTLOSQNYQ 1705
QY 872 KLADENK---SOKLSASDIASSLNKLGVSMALFGTANSVSSNALQIPITQ 917
DB 1706 DADDAKRTAYSQAVNAA--ATILNKTAGG-----NTPKADVERAMQAVTQ 1748

RESULT 23

US-09-815-242-12610
; Sequence 12610, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078

;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: PaetSEQ for Windows Version 4.0
;; SEQ ID NO 12610
;; LENGTH: 5795
;; TYPE: PRT
;; ORGANISM: Staphylococcus aureus
US-09-815-242-12610

Query Match 4.2%; Score 196; DB 9; Length 5795;
Best Local Similarity 18.6%; Pred. No. 0.0005;
Matches 221; Conservative 165; Mismatches 431; Indels 372; Gaps 54;
QY 1 MSNINVIKSNIQAGINSTKSGKXLYLAI PKDYDPQKGTGLNDFIKAADELGIARLAEAP 60
DB 2937 LEDIQVTQNGQA-LNNAMKGLRD-----SIANE-----TTVKASQNYTDA-----SP 2978
QY 61 NHTETAKKSDVT-----VNOFLSLTQTGIAISATKLEKFLQKHSNKLAKGLSDSVENIDRK 116
DB 2979 NNQSTYNSAVSNGKIINQTNPTMTDSAITQATTQ-----VNNAKNGLNGAENLRNA 3031

QY 117 LKASNVLSLSTSLFGLTALAGIELDSLKKGDAPDALAKASIDLINIEIGNLSQS----- 172
DB 3032 QNTAKQNLNTLSHLTNQKSAIS-SQIDRAGHVSVEVTAAKNAATLNTQMGNEQAITHDQ 3090
QY 173 ---TQTTEAFSSOLAKGLSTTSQAKGFSNIGKLNQNLNFSKTNL--GLEIITGLLSGISA 227
DB 3091 NTVQGVNFTDADKAKRDYANVSRAETILNKTKGTQTSKQDVEAIIQNTSNAKLN- 3149
QY 228 GFALADKNASTGKVAAGFELSNOVIGNVT-----KAISYVLAQRVAAG---LSTTG 277
DB 3150 ---GDQNVTVNAKNTA-----KHALNLTISNNAQKEDLTTKIDQATTVAGVEAVSNTG 3199
QY 278 AVAALITSSIMLAISPLAFMAADKFNANALDEPAKQFRKFGVDGDHLLAEYQGVGT- 336
DB 3200 TQLNTAMANLQNGINKANTILASENYHDADSDKKTAA--YTQAVTNAENIL---NKNSGSN 3254
QY 337 -----IEASLTTISTALGAVSAGVSAAVSAVGTPTALLVAGVTGLISGILEASKQAM- 390
DB 3255 LDKAAVENALSQVNTAKGALNGHNLEQAKSNANTTTINGLOHLTTAOKDKLKQVQOQAQN 3314
QY 391 -----FESVANRLQG-----KILEWEKQNGQNYPD-----KGYDSRYAAYL 427
DB 3315 VAGVDTVKSSANTLNGAMGLRNSIQDNTATNGQNYLDATESNKTYNNNAVDANGVIN 3374
QY 428 A-----NNLKPLSELNKE----- 440
DB 3375 ATSNPNMDANAINQIATQVTSKKNALDGHNTLNTQAKQTATNAIDGATNLNKAQKDALKAQ 3434
QY 441 -LEAERVTAITQ-QRWDNNGIGELAGITKLG-----ERIKSGKAYADAPEDGKV----- 487
DB 3435 VTSQORVANVTISIQQTANELNTAMQLQHGIDDENATKQTKYRDA-EQSKKTYADQAVA 3493
QY 488 -----EAGSN-----ITLDAKTGIIIDISNSNGKKTQAL-----HFTSPL 521
DB 3494 AAKAILNKQTSNSDKAAVDRALQVSTKDALNGDAKLAEAKAQAQNLGTLNHIITNAQ 3553
QY 522 LTA-----GTSRRLTNGKYSYINKL-----KFRGVKNQVTTDGAASSKLD 563
DB 3554 RTALEGOINQATTVDGVNTVKTNANTLDGAMNSLQGSINDKATLNRNQNYLDADESKRNA 3613

QY 564 FSKVIQORVAETEGTDEIGLIVNAKAGNDDIFVGGQKKNIDGGDGHDRFVYSKD--GGFGN 621
DB 3614 YT---QAVTAAEG-----ILNKQTGNN-----TSKADVD--NALNTVTRAKALNGAEN 3657
QY 622 ITVDGTSATEAGSYTN-----RKVARGDIYHEVVKRQ-----ETKVGKR 661
DB 3658 LRNTKTSATN-----TINGLFLNLTQLOKDNLKHQVEQAQNVAGVGVKDKGNTLNTAMGAL 3713
QY 662 TETIQ-----YRD-YELRKVGY----- 677
DB 3714 RTSIQNDNTTKTSQNYLDASDSKNNTYNTAVNNANGVINVTNNPNMDANANGMANQVNT 3773
QY 678 -----GYQ-----STDNLKSVVEVIGSQFN-----DVPKGSKFN 706
DB 3774 TKAALNGAQNLAQAKTTAKNNLANLTSINNAQKDALTRSIDGATTVAGVNVQETAKATAELN 3833
QY 707 DIFHSBEGDLDLGGAGDRLFGCKGNDRLSGBEGDLDLGGSG-----DDVLNGGAGND 761
DB 3834 NAMHS-----LQNGI-----NDETQTKTQKYLDAFPSKKSAYDQAVNAAKA-- 3875
QY 762 VYIFRKGDNDTLYDGTGNDKLAFAFADANISDIMIERKEGIIVKRNDHSGSINIPRWYIT 821
DB 3876 --ILTKASQNV-----DKAAVEQA-LQNVNSTKTALNGDAKLNKAAKAKQTGLTTLT 3925
QY 822 SNLQYQSNKTDHKIEQIGKDGSIYTS-----DQIDKILQPKDGTVITSQELK 871
DB 3926 -HINNAQRTALDNEITQATNVGVNTVAKAAQQLDGAQOLETSIRDK--DTTLOSQNYQ 3982
QY 872 KLADENK---SOKLSASDIASSLNKLVGSMALFGTANSVSSNALQPIQ 917
DB 3983 DADDAKRTAYSOAVNAA--ATILNKTAGG---NTPKADVERAMQAVTQ 4025

RESULT 24

US-10-369-493-19792
; Sequence 19792, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19792
; LENGTH: 262
; TYPE: PRT
; ORGANISM: No. US20030233675A1toc punctiforme
US-10-369-493-19792

Query Match 4.1%; Score 191; DB 12; Length 262;
Best Local Similarity 26.2%; Pred. No. 1e-05;
Matches 70; Conservative 35; Mismatches 80; Indels 82; Gaps 13;
QY 532 LTNGKYSYINKLFRGVKNQVTDGE---ASSKLDFSKVIQORVAETEGTDEIGLIVNAKA 588
DB 16 LTGVKPTFLFGSGSSNDNTVQPGKTFPAGDADF-----VEGTEN-----TIQA 62
QY 589 GND--IFVGGQKKNIDGGDGHDRFVYSKDGFGFNITVDGTSATEAGSYTNRKVARGDI 646
DB 63 GNDDDTVLVGSNS-SVSAGDGNQIVIGONGAAQNISADG-----GNGDD 106
QY 647 YHEVVKRQETKVGKRTETIQRYDELKRVGYGQSTDNLKSVEEVIGSQFND---VFKGS 703
DB 107 VITVEA-----NGNNL-----FGAEGNDTLTVIEGS 134


```

QY      704  KFNDFPHSGEGDLDLGGAGBDRFRFGKGKNDRLSGDEGDDLDDGGSGDDVNLGGAGNDVY  763
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      135  R-QSLF-GGLGNDLTUTSGNSNNRLYGGSGDKIPSTNDSLFGGCGDDVLFAGQG-----  187

QY      764  IFRKGGGNDTLYDGTGNDKLAFAIANI  790
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      188  -----GSNRLSGGAGADQFWIANASL  208

RESULT 25
US-09-797-862-32
; Sequence 32, Application US/09797862
; Patent No. US20020102276A1
; GENERAL INFORMATION:
; APPLICANT: PEAK, IAN RICHARD ANGELM
; APPLICANT: JENNINGS, MICHAEL PAUL
; APPLICANT: MOXON, E. RICHARD
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0134
; CURRENT APPLICATION NUMBER: US/09/797,862
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 1098
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-797-862-32

```

| | | | |
|---|------|---|------|
| Qy | 469 | --BRIKSGKAYADAPEDGKV--EAGSNITL---DAKTCIIDI---SNSNGKKTQ---- | 513 |
| Db | 499 | SEQEVKAGDKV--TFRAGKRLKVKQEGANFTYSLQDALTGLTSITLGTGNGAKTEINKD | 556 |
| Qy | 514 | -----ALHFTSPLLTAGTESRERLNNGKVSYNKILKFRGVKQWQVT--- | 554 |
| Db | 557 | GLTITPANGAGANNANTI SVTKDGISAGQGSQKVVVSG-----LKKFGDANFDPITSSA | 610 |
| Qy | 555 | DGEASSKLDFSKVIORVAETEGDETGLIV---NAKAGNDDI-----FVQGGKMNIDGGDG | 607 |
| Db | 611 | DNLTQKNDAYKGLTNLDE-KGTDKQTPVADVNTAATVGLDRLGLGWVISADKTTGGSTEY | 669 |
| Qy | 608 | HDVRFVSKGGF--GN-ITVDGTSATEAGSYVNRKVARGDIYHEVVKQETKV-----GK | 660 |
| Db | 670 | HDQVRNANEVKEKSGNGINVSG--KTVNGRREITPELAKG-----EYVKSNEFTVKETNGK | 723 |
| Qy | 661 | RTETIQRYRDELKVGVGYSTDNLKSVE-----EVIQSOFNDVFKGSKFNDIFHSG | 712 |
| Db | 724 | ETSLV-----KVGDKYYSKEDIDLTTCQPKLKDGNVAAKYQD--KGGKVVSVTDNT | 773 |
| Qy | 713 | EGDDLLDGAGDRLFCGKGNDRLSGDEGDDLLDGSGDDVLNGAGGANDVYIFRK-GDGN | 771 |
| Db | 774 | EA-----TITKSGSVYTCQVADAI-AKSGFEL-----GLADEADAKRAFDDKT | 817 |
| Qy | 772 | DTLYDGT-----GNDKLAF-----DANISIMERTKEGLIIVKRNDSHSGSINIPRWITSN | 823 |
| Db | 818 | KALSAGTTEIVNAHDKVRFANGLNTKVSAAVVEST-----DANGD-KVTTTFVKTD | 867 |
| Qy | 824 | L-----QNYOSNKTDHKIQLICKQG-----SYITSDDID---KILQD | 858 |
| Db | 868 | VELPLTQIYNTDANGKKITKV--KQGTWKYELNADGTADMTKEVTLGNVDGCKKVVKD | 926 |
| Qy | 859 | -----KKDGT-----VITSOELKKLADENKSQK-----LSA- | 884 |
| Db | 927 | NDGKWHAKADGTADTKTGEVSNDKYSTDEKHVVSLDPNDQSKGKGVVIDNVANGDISAT | 986 |
| Qy | 885 | -----SPIASSLNKL-----VGSNALFGTANSVSSNALQIPITOPTQG | 921 |
| Db | 987 | STDAINGSQLYAVAKGVITNLAGOVNLEGKVNKVKGRADAGTASALAASQLPQATMPGKS | 1046 |
| Qy | 922 | ILA 924 | |
| Db | 1047 | MVA 1049 | |
| RESULT 26 | | | |
| US-10-185-990-11 | | | |
| ; Sequence 11, Application US/10185990 | | | |
| ; Publication No. US20030073109A1 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Pan, Jae-Gu | | | |
| ; TITLE OF INVENTION: JAE GU PAN ET AL | | | |
| ; FILE REFERENCE: 02589.000100 | | | |
| ; CURRENT APPLICATION NUMBER: US/10/185,990 | | | |
| ; CURRENT FILING DATE: 2002-06-28 | | | |
| ; NUMBER OF SEQ ID NOS: 11 | | | |
| ; SOFTWARE: PatentIn version 3.1 | | | |
| ; SEQ ID NO 11 | | | |
| ; LENGTH: 1626 | | | |
| ; TYPE: PRT | | | |
| ; ORGANISM: Bacillus subtilis | | | |
| US-10-185-990-11 | | | |
| Query Match 4.0%; Score 188; DB 15; Length 1626; | | | |
| Best Local Similarity 20.2%; Pred. No. 0.00027; | | | |
| Matches 185; Conservative 112; Mismatches 333; Indels 286; Gaps 39; | | | |
| Qy | 63 | TETAKSVDVTVNOFSLTQTGIAISATKLEFKLQKHSTNKLAKGLDSVENIDRKLKGSN | 122 |
| Db | 785 | TQTAQERSDLVTGYGSTSTAGYASS-----LIAGYGSTQ-----TAG | 821 |
| Qy | 123 | VLSTLSSFLTALAGIELDSLIKKGAAPALAKASIDLI-----NEITGNLSQ----- | 171 |

Db 822 YESTLTAGYSTQAOB--NSSLTGYGSTSTAGPAS--SLIAGYGSTQAGYKSTLTAGYG 879
Qy 172 STOTIEAFSSOLAKLGSTISOAGFSNI-----GNKLQN--LNFSTKNLGLIITGLLSGIS 226
Db 880 STQTAEGSSLTGYGSTATAGODSSLIAGYSSLTSGIRSFALTAGYGSTLIAGLSVLI 939
Qy 227 AGFAL-----ADKNASTGKVAAGFELSNOVIGNVTKAISSYVLAORVAA 271
Db 940 AGYSSLTSGIRSTLTAGYSGNQIASYSSLIAGHE--SIQVAGNKSMLIAGKSSO--TA 996
Qy 272 GLSTTGAAVALITSSIMLAISPLAFMAADKFNHANALDEFAKQFRKFGYDGHLLAEYQ 331
Db 997 GFRSTLIAGA--GSVQLAGDRSRLIAGADS-----NOTAGDRSKLAGNNSYLTAGD 1046
Qy 332 RGVGTIEASITTTISALGASVAG---VSAAAVSAGVTPIALLVAGV--TGLISGILEAS 386
Db 1047 RSKLTGCHDCTLMAGDOSRLTAGKNSVLTAGARSKLIGSEGSTLSAGEDSTLIFRLWDGK 1106
Qy 387 --KQAMPESVANRLQKGKIL-----EWEKQNGQNYFDRGYDSRYAAYL 427
Db 1107 RYQLVARTGENGVADIPYVNEDDDDIVDKPDEDDDIIEVKPMGVFD----- 1155
Qy 428 ANNKLFLSELNKL--BAERVIAITQORWDNN-----IGELAGITKL 467
Db 1156 YKNLG--TEASKTLFADATATLTYHNLNDFPAVGYYQOHLGLGLPATLVGALLGST-- 1211
Qy 468 GERIKSKAYADAFEDGKKVEAGSNITLDKATGIIDISNSG---KKTQALHTSPPLTA 524
Db 1212 -----DSQVIGIPGIPMNPDSSEKALDAVHAAGWTPISASALGYGKVDAR 1256
Qy 525 GTESRRLTKGYSYINKLFGFRVKMWQVTDGEASSKLDKFSKVIOVA-----ET 574
Db 1257 GT-----PFGEKAGYTTAQAEVLGKYDAGKLLLEIGIGFRGTSGPRE 1299
Qy 575 EGTDEIGLIVN--AKAGNDDI--FVQG-----GMMNIDGGD---GHRV 611
Db 1300 LITDSIGDLVSDLLAALGPDKYAKNYAGAEFGGLTKTVADYAGAHGLSGKDVLSGHS- 1358
Qy 612 FYSKGGFG-NITVDGTSATEAGSYTNVRKARGDIIYHEVVKQETKVGKRTETIQYRD- 669
Db 1359 -----GGLAVNSMADLSTSKWAGPYK-----DANYLAYASPTQSGDKVLNIGYEND 1405
Qy 670 -----YELRKVG--GYOS--TDNLKSVEEVIGSOFNDVFKGSKENDI----- 708
Db 1406 PVFRALDGSFTNLSGLVHDKAHBESTDNIVS-----FNHYASTLWNLVLPFSTANL 1457
Qy 709 -----FHSGEADDL--LDGG-----AG 723
Db 1458 STWVSHLPSAYDGMTRVLESFPYEQMTRDSTIIIVANLSDPARANTWVQDLNRNABPTHG 1517
Qy 724 DDLRFGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDNNTLYDGTGNDKL 783
Db 1518 NTFIIGSDGNDLIQGGKGADEIEGKGNDTRDNSGHTFLFSGHFGQDRIIGYQPTDRL 1577
Qy 784 AFADANISDIWIERTK 799
Db 1578 VFQAGDSTDLRDAK 1593

RESULT 27

US-10-311-879-28
; Sequence 28, Application US/10311879
; Publication No. US20030186275A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; TITLE OF INVENTION: Antigenic Peptides
; FILE REFERENCE: toxin
; CURRENT APPLICATION NUMBER: US/10/311,879
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 2659
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-10-311-879-28
Query Match 4.0%; Score 186.5; DB 12; Length 2659;
Best Local Similarity 19.0%; Pred. No. 0.00073;
Matches 186; Conservative 129; Mismatches 353; Indels 309; Gaps 44;
Qy 4 INVIKSIQAGLNSTKSLKNL-----YLAIPKDY---DPQKGGTLNDFIKADELGIAR 55
Db 1784 VNTIKQN--ATNLNSAMGNLRQAVADKQVKTREYADADATKQNAVNSAVSSAETI----- 1838
Qy 56 LABEENHTETAKSVDTVNFQFLSLTQTGIAISATK-----LEKFLQKHSNKLAKGLDSV 110
Db 1839 -----INQTNPTMSVDDVRATG-----AVTSNKVALNGYEKLAQ--SKTDAARALDA 1886
Qy 111 ENIDR-----KLGKASNV-----LSTLSPFLGTALAGIE-----LDS 142
Db 1887 PHLNNAQKADVKSKINAASNIAGVNTVKQGGTDLTAMNLOGAINDEQTTLSNQYQDA 1946
Qy 143 LIKKGDAAPDALAKASIDILNLIIGNLSQSTOTIEAF---SSOLAKLGS--ISOAK--G 195
Db 1947 TPSKKTAYTNAV--QAAKDILKNSGQNKTKDQVTEAMNQVNSAKNNLDGTRLLDQAKTA 2005
Qy 196 FSNIGNKLQNLNPSKTNLGLIITLGLSGISAGFALADKNASTGKKVAAGFELSNOVIGN 255
Db 2006 KQQLNNWTHLTTAOKTNLTNQINSGT---TVAGQVTVQSNANT---LDQAMWTLRQSIAN 2059
Qy 256 --VTKAISSVLA--ORVAAGLSTTGAAVALITSSIMLAISPLAFMAADKFNHANALDE 311
Db 2060 KDATKASEDYVDANNDKQTYANNAVAAAEIINANSNPEMNPSTITQKABQVNSS----- 2114
Qy 312 FAKQFRKFGYDGHLLAEYQVGVGTIEASLTITSTALGASVAGSAAVSGVATGPFIALL 371
Db 2115 -----KTALNGDENLAAAKQNAKTYLNTLTSITD----- 2143
Qy 372 VAGVTGLISGILEASKOAMFESVANRLQKILEWEKONGQNYFDKGYDSRYAAYLANNL 431
Db 2144 -AOKNLLISQITSATRVSGVDIV-----KQNA--QHLDOA-----MASLONG- 2182
Qy 432 KFLSELNKELEABERVIAITQORWDNNIGELAGITKGERIKSKAYADAFEDGKKVEAGS 491
Db 2183 -----INNE-----SQVKSSEKYRDA--DTNKQOEYDN 2208
Qy 492 NIT-----LDAKTGIIDISNS-----NGKKTQALHTSPPLTAGTESRRLTKGYS 538
Db 2209 AITAAKAILNKSTGPNPTAQNAVEAALQRVNNAK--DALNGDAKLIAAQNAKQHL--GTLT 2265
Qy 539 YINKLFGFRVKMWQVTDGEASSKLDKFSKVIOQVVAETEGTDEIGLIVNAKAGNDDIFVQGG 598
Db 2266 HIT-----TAQRNDLTNQISOATNLGYESV----- 2291
Qy 599 KXNIDGGDGHDRVYFYSKGGFGNI---TVDGTSATEAGSYTNVRKARGDIIYHEVVKRQET 656
Db 2292 KQAN-----SLDGAAGNLQTAINDKSGTSLASQNFDADEQKRNAYNOAVSAET 2341
Qy 657 KVGKRT-----ETIQYRDYELRKVGYGYSQTDNLKSVEEVIGSOFNDVFKGSKENDIFH 710
Db 2342 ILNKQTPNTAKTAVQEQALNNVNNAKHALNGTQNLNNAKQAIAITAIN----- 2388
Qy 711 SGEGDDLLDGGAGDRLFGCKGNDRLSGDBGDDLLDGGSGDDVLNGGAGNDVYIFRKGDG 770
Db 2389 -----GASD---LNQKQDKALKAQ-----ANGAQ-----RVNSA 2414
Qy 771 NDTLYDGT-----GNDKLAFAD-----ANISDIMERTEKEGIIYKRNDSHSGSI--NIPR 817
Db 2415 QDVQHNATELNTAMGTLKHAIADKTNTLASSKYVNADSTKQNAVYTKVTVAEHIISGPT 2474
Qy 818 WYITSNLQNYQSNKTDHKEQLIGKDGYSITSDQIDKILQDKGDTVITSQELKULADEN 877
Db 2475 VWTTPSEVTAANQVNSAKQELNG-----DERLEAKQONANTAIIDALTQL---N 2520
Qy 878 KSQKLSASDIASSLNKL 894


```

; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12996
; LENGTH: 6281
; TYPE: PR1
; ORGANISM: Staphylococcus aureus
; US-09-815-242-12996

Query Match      4.0%; Score 186.5; DB 9; Length 6281;
Best Local Similarity 18.6%; Pred. No. 0.0027;
Matches 224; Conservative 153; Mismatches 477; Indels 349; Gaps 51;

Qy 1 MSNIVKSNIOAGLNT--KSGLNLYLAIPKDYDPQKGTDLNDFTKADELGIARLAE 58
Db 208 IQNVSSAKNALGQNVTAQNAKNAKLNLSINNAQKRDLTTKIDQATTVAGVEAVSN 267
Qy 59 EPNHTEAKSV-----DTNQFLSLT-----QTGIALSATKLEKFLQKSTNKLAK 105
Db 268 TSTQLNTAMANLQNGINDKNTLASENYHDADSKTAYTQAVTNNAENILNKSGSNLDK 327
Qy 106 GLDSVENIDRKLKASNVLS-----TLSSPLGTALAGI-----ELDSLKKGDAAPDA 153
Db 328 --TAVENALSOVANAKALNGNHNLEQAKSNANTINGLOHLTTAQKDKLQKQVQOQNV 385
Qy 154 ----LAKASIDLNEIIGNLQSQTITIEAFSSQLAKLGSTISQAKGFSNTGNKLQNLN-- 207
Db 386 AGVDTVKSSANTLNGAMGTLRNSIQDNTATKNGQNYLDAT---ERNKNTYNNNAVDSANGV 442
Qy 208 ---PSKTNLGLIITGLLGSISAGFALADKNASTGKKVAAGFELSNOQVIGNVTKAISYV 264
Db 443 INATSNPNMDANAINQIATQVT-----STKNALDG---THNLTOAQQTATNAIDGATNLN 494
Qy 265 LAQRVAAGLSTGAVAAALITSSIMLAISPL-----AFWNAADKFNHANALDEF--AKQFR 317
Db 495 KAKQDKALKAQVTSQRVANVTISQOATNELNTANGLOHQHGDIDENATKQTKYEDASQSK 554
Qy 318 KFGYD----GDHLAEYQGVGTIEAS-----LTTISTAL--CAVSAGVSAAVGSAVG 365
Db 555 KTAVDQVAAAKAILNKQTSNSDKAIVDRALQOVTSTKDALNGDAKLAELAKAAKQNLG 614
Qy 366 TPIALLVAGVTGL-----ISGILEASQA--MPESVANRLOGKILEWEKONGQNYF 415
Db 615 TLNHTIWAQRTDLEGINQATTVGVNTVKTNANTLDGAMNSLQGSINDKATLURNQYL 674
Qy 416 DKGYDSRYAAYL-----ANNKFLSE 436
Db 675 DAD--ESKRNATQAVTAEGILNKQTCGNTSKADVNDNALNAVTRAKAALNCADNLR---- 729
Qy 437 LNKELEAERVI-----AITQQRWDN-----NIGELAGITIKLGERI----- 471
Db 730 -NAKTSATNTIDGLPNLTQKQNLKHKQVEQAQNVAGVGVKDKGNTLNTAMGALRTSIQ 788

472 ----KSGYADAPEDGKKEAGSNITLDAKTIIDISNS-----NGKKTQ----- 513
789 NDNTTKTSQNYLDASDNK---NYNTAVNANGVINATNPNMDANALNGMANQVNTTK 845
514 -ALHFTSPLLTAGTESRERLTNG-----KYSYINKLKFGVRKYNQVTDGESSK 561
846 AALNGAQNLAQAKTNATNTINNAHDLNKQKQKDALKTQVNAQVRSDANNVQHTATELNSA 905
562 LDPSKVIQVRAETEGTDEIGLIVNA-----KAGN-----DIFVIG 596
906 MTALKA--AIADKERTKASGNYVNAQDEKQRAYDSKYVTAENIISGTPNATLTVNDVNSA 963
597 QGRANID---GGDGHDRVFSKDGCGFNI-----TVDTGTSAT 630
964 ASQVNAAKTALNGDNNLRV--AKEHANNTIDGLAQLNNAQKAKLKEQVQSATTLDGVTQTV 1021
631 EAGSYTVN--RKVARGDIYHEVV-----KROE----- 655
1022 KNSSTLNTAMKGLRDSIANEATIKAGQNVYTDASPNNRNEYDSAVTAAKAIINQTSNPTM 1081
656 ----TKVGRTEITQYRDYELRKVGYGYQST--DNLSKSVEEVIGSQFN----- 697
1082 EPNITQVTSQVTTKEQALNGARNLAQAKTAKNNLNLSINNAOKDALTRSIDGATTV 1141
698 ----DVFKSGKFNDIFHSGEGDLDLGGAGDDRLFGKGNDRLSGDEGDDLDG----- 747
1142 AGVNOETAKATELNNAMHS-----LQNGI-----NDETQTKQTKYLDABPSK 1185
748 GSGDDVLANGAGNDVYIFRKGDNQDNTLYDGTGNDKLAFAFADANISDIMERTKEGIIVKRN 807
1186 SAYDQAVNAKA---ILTKASQNV-----DKAAVEQA--LQNVNSTKTALNGDAKLN 1233
808 DHSGSINIPRWYITSLNQNYSNKTQDHKEIQLIGKQGSYITS-----DQIDKILQ 857
1234 EAKAAAKQTLGLT--HINNAQRTALDNEITQATNVEGVNTVKAKAOLDGAMGQLETSIR 1292
858 DKKDGVTITSQELKLADENK---SOKLSASDIASSINKLVGSMALFGTANSVSSNALQP 914
1293 DK--DTTLOSQVQDADDAKRTAYSQAVNAA--ATILNKTAGG---NTPKADVERAMQA 1344
915 ITQ 917
1345 VTQ 1347

RESULT 30
US-10-369-493-19046
; Sequence 19046, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(S2052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19046
; LENGTH: 1136
; TYPE: PR1
; ORGANISM: Anabaena PCC7120
; US-10-369-493-19046

Query Match      3.9%; Score 180.5; DB 12; Length 1136;
Best Local Similarity 19.6%; Pred. No. 0.00054;
Matches 197; Conservative 125; Mismatches 354; Indels 327; Gaps 44;
```



```

Query Match      3.9%; Score 180; DB 12; Length 273;
Best Local Similarity 25.8%; Pred. No. 6.7e-05;
Matches 88; Conservative 37; Mismatches 100; Indels 116; Gaps 19;

Qy 459 GELAG--ITK-LGERIKSKAYADAFEDGKKVEAGSNITLDAKTIIDISNSNGKKTQAL 515
Db 30  GDLGRAITEALGRCPCSSLYXRSI--DFEKTNGQATVLDQS-----VGNTQ----- 77

Qy 516 HFTSPLTGTAGTESRERLTNGKYSYINKLFGKRVKNQVMTGGEASSKLDPSKVIQVVAETE 575
Db 78  YFTA-----AGFDR-----TVGLTTL--DLQFGSTNSDDVT-----LKPNTLTFAGD 118

Qy 576 GTDEI-----GLIVNAKAGNDDIFVGCQKMNIDGGGHDHVRFY-----SKDGGFGNI 622
Db 119 GADPVEGTGKNTVITGNGEDTVLVGSGS--SVSTGDNQDQFVIGVNSPASVTSADGGNGN- 176

Qy 623 TVDGTSAEA--GSYTVNRKVRGDIYHEVVKRQETKVGKRTETIQRDYELRKYGVGYQS 681
Db 177 --DEVTVVEANGSNL-----PGGAG 195

Qy 682 TDNLKSVEEVIGSFQFNDVFGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEG 741
Db 196 ADTLTVVE-----GSR--QLSFGSGGNDTLKSNNGSNRLYGGSGDDKLFSSVN 241

Qy 742 DDLDDGGSGDDVLNGGAGNDVYIFRKGDNGLDLYDGTGNDK 782
Db 242 DSLF-GGDDDDVLPAQO-----QGRLTGGTGADQ 271

RESULT 33
US-09-841-786-6
; Sequence 6, Application US/09841786
; Patent No. US20020054883A1
; GENERAL INFORMATION:
; APPLICANT: NAGARAJA, T. G.
; APPLICANT: STEWART, GEORGE C.
; APPLICANT: NARAYANAN, SANJEEV K.
; APPLICANT: CHENGAPPA, M. M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
; FILE OF INVENTION: VACCINE AND PREPARATION THEREOF
; FILE REFERENCE: 30296
; CURRENT APPLICATION NUMBER: US/09/841,786
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 09/558,257
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 773
; TYPE: PRT
; ORGANISM: Fusobacterium necrophorum
US-09-841-786-6

Query Match      3.9%; Score 179; DB 9; Length 773;
Best Local Similarity 21.4%; Pred. No. 0.00039;
Matches 172; Conservative 114; Mismatches 310; Indels 208; Gaps 38;

Qy 111 ENIDRKLKGSNVLT-----LSSFLGT-ALAGIELDSLKKGDAAPALAKAS--ID 160
Db 14  KNVDITTEDKNITSTGGLGTAGLASAGTVAVTNIKRNS---GVTVENSFVKAEEKVN 69

Qy 161 LINEIIGNLSOSTQTIEAFSSQLAKLGSTISQAK-----GFSNI---GNKLNLF----- 208
Db 70  VRSDITGNVA-----LTAYQGPVGAIGTGAAYAEINLSNGRNSISIKNKLKGNKIDVIVK 124

Qy 209 SKTNLGLIEITGLLSIGSAGPALADK-----NASTGKKVAAGPELS 249
Db 125 DKSELRAE--AKGLTVGAAGAIIISKAKVENWSEVEIEKSTFNEENRVTSPSKIGIREI- 182

Qy 250 NOVIGNVTAKISVYLAQVAGLSTTGAVA--ALITSSIMLAISPLAFMNAADKFN----- 304
Db 183 -----NVKVENRVTAESQGA---SVCVAVAGAGIIESEAKDAGS--SYLKVSTKSGRSIP 232

```

```

Qy 305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTITISTALGAVSAG-----VSAAA 359
Db 233 HADNVNWEATHKQKVTAVSKAVTSGVLGGVTKABATAAGKTMTVEVEGNLFRTRLNA 292

Qy 360 VGSAGV-----TPIALLVAGVTGLI--SGILEASKQAMPESVANRLOQKILEWEKONGG 411
Db 293 ISKVEGLDEDKVTAKSVSVSGNGGAGAGAVNTSTAQSNTEV--RLR-----KQDYE 344

Qy 412 QNYFPGKYSRVAAYLANNLKFLSELNKELEAERVIAITQORWNNIGEL--AGITKLGE 469
Db 345 NNDYTK-----KYISEVN-----ALALNDTKNEANIESLAVAGVHAQG-- 382

Qy 470 RIKSGKAYADAFEDGKKVEAGSNIT--LDAKTIIDISNSNGKKTQALHFTSPLTAGTES 528
Db 383 ---TNKAFTRSNTLSTTVNGGNSQLRAKALAKENYGNVKT-----GGALVGAETAA 434

Qy 529 RERLTNGKYSYINKLFGKRVKNQVMTGGEASSKLDPSKVIQVVAETEGTDEIGLIV- 584
Db 435 VENYTK---STTGALVAG--NWEIGD---KLETIARDNTITVRVNGDGTGGLVGKNG 483

Qy 585 -----NAKAGNDD--IFVGCQKMNIDGGGHDHVRFYSGKGGFGNITVDGTSATEAG 633
Db 484 ISVKNITISGETKSSIEDKARIVGTGSVNVVDALNELDVLQKSGGGYGGIGVNV----- 537

Qy 634 SYTVNRKVRGDIYHEVVKRQETKVGKRTETIQRDYELRKYGVGYQSTDNLKSVEBEVIG 693
Db 538 -----DYNVVIKKNVEAKIGRH-----AIVETTG 561

Qy 694 SQFNDVFGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGD--EGDDLLDGGSGDD 752
Db 562 QOEYQAFTRAKNVIL-----GKGDAAAAAAISNVHISNEMDIKNLAKQVASSQ 609

Qy 753 VLNGGAGNDVYIFRKGDNGLDLYDGTGNDKLAFAFADANIS--DIMIERTKEGIIVKRNDHSG 811
Db 610 LITKSNKNITTLASSSESNNVNH--GVAEARGAGAKATVSVKNOINRTNNVDLAGKIKTEG 668

Qy 812 SINIPRWYITSNLQYQSNKTDHK 835
Db 669 NINVYAGYD-----KYNISIKTNSK 688

RESULT 34
US-10-193-764-57
; Sequence 57, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; FILE OF INVENTION: MOLECULAR WEIGHT PROTEINS
; FILE REFERENCE: 1038-1239MIS
; CURRENT APPLICATION NUMBER: US/10/193,764
; PRIOR FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 57
; LENGTH: 992
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-193-764-57

Query Match      3.9%; Score 179; DB 12; Length 992;
Best Local Similarity 22.2%; Pred. No. 0.00056;
Matches 207; Conservative 110; Mismatches 329; Indels 286; Gaps 54;

Qy 136 AGIELDSLKKGDA---APDALAKASIDLINEIIGN--LSQSTOTIEAFSSQLAKLGSTIS 191
Db 16  AGI--DSEFFPGSGTGTSKPTNGEQPVLNTNETISNYLKSJGTVMNITAKKNLTVNSSI- 72

Qy 192 QAKGFSNIGNKLNLFNS--KTNLGLIEI---IT---GLLSGISAGFALADKNASTG----- 239

```



```

Qy 881 KLSASD---IASSLNKLVGSMALFCTANSVSN 910
Db 793 TLTAKGSIAGSINAANVTTLTTGLTTVEGS 824

RESULT 36
US-09-815-242-5835
; Sequence 5835, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zvekind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078.
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5835
; LENGTH: 2434
; TYPE: PRN
; ORGANISM: Staphylococcus aureus
US-09-815-242-5835

Query Match 3.8%; Score 177.5; DB 9; Length 2434;
Best Local Similarity 17.9%; Pred. No. 0.0028;
Matches 229; Conservative 164; Mismatches 450; Indels 433; Gaps 51;

Qy 2 SNINVIKSNTOAGLN---STKSGLNLYLAIPKDYDPQKGGTLDNDFKAADEL-----51
Db 1003 SELTAMSNLQNGINDEAATKAAQY-----TDADREKQTAYNDVAVTAARTLLDKTAGS 1056
Qy 52 -----GIARLAEEPNTHTAKSVDTVN-----QFLSLTQ---81
Db 1057 NDNKAAVEQALQRVNTAKTALNGDERLNEAKN---TAKQVATMWSHLTDAQKAWLTQSIE 1113
Qy 82 -----TGIAISATKLEFL-----QRHSTNKL-----AK 105
Db 1114 SGTTVAGVGGIQAAGTLDQAMNQLRQSIASKDATKSSDYQDANADLQNAVNDVATNAE 1173
Qy 106 GLDSVEN-----IDKLGKASNVLSLS-----SFLG---TALAGIELDS 142
Db 1174 GIISATNPNMNPPTINQKASQVNSAKSALNGDEKLAATAKQTAKSIDIGRLTDLNNAQRTA 1233
Qy 143 LIKKGDAAPD---ALAKASIDLNEIIGNI-----SQTQTIEAFSSQLAKGS 188
Db 1234 ANAEVDQAPNLAUAVTAAKNKAATSNLTAMGNLHUALAEKNDKTVKSVNVTADQPPQQAAYDT 1293
Qy 189 TISQAKGFSNIGNKLNLFSTKTLGLEIITGLLSGISAGFALADKNASTGKVKVAAGFEL 248
Db 1294 AVTQAEAITNANG---SNANETQVQAALNQLNAQKNDLNGDKNVQAQAKESAKRALASYNL 1351

```

RESULT 37

```

US-09-971-536-68
; Sequence 68, Application US/09971536
; Patent No. US20020159976A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka

```

```

Qy 249 SN-----QVIGNVTKAISSYVLAQRAAGLSTTGAVAAALITSSIMLAISPLAFMAAD 301
Db 1352 NNAQSTAATSQIDNAT--TVAGVTAAQNTANELN-----AMGOLQN 1391
Qy 302 KFNHANALDE-----FAKQFRKEGYDGDHLLAEYQRGV-----GTIBASLTIS 345
Db 1392 GINDQNTVKQOVNFTDADQKGDAYTNVNTAQ---GILDKAHGQNTWTKAQVEAALNQVT 1448
Qy 346 TALGAVSAGVSAASAAVGSVAGTPIALL-----VAGVTGL--ISGILEASKQA 389
Db 1449 TAKNALNGDANVRQAQSAKANLGTLLHNLNAQKODLTSQIEGATTVNGVNGV--KTKAQ 1506
Qy 390 MPESVANRLQCKILEWEKQNGQYFDKGYDSRYA-----AVL-----ANNLKFLS 435
Db 1507 DLDGAMQRLQSAJANKDQTKASENYIDADPTKTAFDNITQAESYLNKDHGANKQKQAV 1566
Qy 436 E-----LNKELEAEARVIAITQORWMDN-----NIGE-LAGITK 466
Db 1567 EQAISOVSTSTENALGNDANLQRAKTEAIAQIDNLTLHTLTPQKTAALKQOVNAAQRVSGVTD 1626
Qy 467 LGER-----IKSGK-----AYADAFEDGKKVEAGS-NIT 494
Db 1627 LKNSATSLNANMDQLKQAIADHDITI VASGNYTNASPDKQAYTDAYNAAKNI VNGSPNVI 1686
Qy 495 LDKATGIIIDISNSGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNQVOT 554
Db 1687 TNAADVTAATQRVNNAET-GLNGDITNLATAKQAKDALLR-----QMT 1727
Qy 555 DGEASSKLDPFSKVIQVRAETEGTDEI-----GLIVNAKAGNDDIFVGGKKNID 603
Db 1728 HLSDAQKQSIGQIDSATQVTGVQSKDNATNLNANQLRNSIANKDDVKAQK--PYVD 1785
Qy 604 GGDGHRDVFYKDGKGFNI-----TVDGTSATEAGSYTVNRKRVARGDIYHEVVRKQET 656
Db 1786 ADRDKQNAVNTAVTNAENIINATSQPTLDPASVTOAANQVSTNKTALNGAQNLANKKQET 1845
Qy 657 -----KVGKRTETIOYRDYELRKYGVGYQSTDN 684
Db 1846 TANINQLSHLNNNAQKODLNTQVTNAPNISTVNGQVTKAEQL-----DQAMERLINGIQDKQ 1902
Qy 685 LKSEVEVIGSQFNDVFKSGKFN-----DIFHSGEGDDLLDGGK-----722
Db 1903 VK---QSVNFTDADPEKQATAYNNAVTAENIINQANGTNANQSQVEAALSTVTTTKQALN 1959
Qy 723 GDRRLFGGKGNDRLSGDEGDDLLDGGSG-----DDVLNGGAGND 761
Db 1960 GDRKVTDAKNANQTLSTLDNLNNAQKQAVTGNINQAHTVAEVTOAIQTAAQLNTAMGN- 2018
Qy 762 VIIFRKDGDNDTLYDGT-GNDKLAFAFADANISDIMIE--RTKEGIIVKRNDSHGSINIPRW 818
Db 2019 -----LKNLSNDK--DTTLGSONFADADPEKKNAYNEAVHNAENLNK-----STGTNVPKD 2068
Qy 819 YITSNLQYQSNKTDHKIEQLICK-----DG-SYITSDQIDKILQPKDGTVITSQ 868
Db 2069 QVEAAMQVNTAKAALNGTQNLKAKQAHANTALDGLSHLTNAQKALKQLVQOSTTVA--2126
Qy 869 ELKKLADENKSKLSASDIASSLNKLVGSMALFCTA-----NSVSSNA 911
Db 2127 -----EAQCNQKANNVDAAMDKLQRSIADNATTKQNYTDSQNKKKDAYNNAVTTA 2179
Qy 912 LQBITQPTQGILAPSV 927
Db 2180 QGIIDQTTSPTLDPVT 2195

```



```
Db 129 NY-----ATKLEHQLEBAVQAKLEAEQYHKIPYEIKTFVTVIDRVYG 176
Qy 119 -----KASNVLSLTSFLGTALAGIELDSLIIKKGDAAAPDALKAKASIDLINEI 165
Db 177 KTRDLRLSRTPKAKAQLRSLDIYITVAMKAREVQDAKAGNLDK---AKAAVDQINQY 233
Qy 166 IGNLSQSQTTEAFSSQLAGSTISQAKGSNIGNKLNLFPSKT-NLGLIITGLLSG 224
Db 234 LPKVT-----DAFKTELTEVAKKALDADE-AALTPKVESVSAINTQKAVELTAVPVNG 286
Qy 225 -----ISAGFALADKNAKSTGKVAAGFELSNOVTNKATSSYVLAQRVAAGLSTTGAV 279
Db 287 TLKQLSAAANEDTVNVTNR-----IYKVDGNIPFALNT-----ADVSLSDTGKT 332
Qy 280 AALITSSIMLAISPLAFNAAD-KFNHANALDEFAKQFRK-----FGVDGDHLLAEYQGV 334
Db 333 ITVDAST-----PFENNTYKVVVKIGDKNGKGFKEDAFTFKLNRDAVVTQV---F 381
Qy 335 GTIEASLTTISTALGA-----VSAGVSAAGVSAVG 365
Db 382 GTNVTNNTSVNLAAGTFDDBTLTVFDFKLAPETVNSSNVTTIDVETGKRPIVIASTSG 441
Qy 366 TPIALLV--AGVTG-----LISGILEASKQAMFESVAN-----RLQKIL 403
Db 442 STIITLKEALVTGQYKLAINNVTLTGYNAEAYELVFTANASAPTATVAPTTLGGTTL 501
Qy 404 E-----WEKONGQON-----YFDKGYDSRYA-----AYLANNLKL-----SE 436
Db 502 STGSLTNVMKLAGGVNEAGTYYPGLQFTTTFATKLDESLADNLFVLVEKESGTVVASE 561
Qy 437 LNKELEAERVTAITQ-QRWDDNIGELAGITKLERIKSGKAYADAFEDGKKEVAGSNITL 495
Db 562 LKYNADAKMVLVPKADLKENTIIYOI-----KIRKG-----LKSDDKGIELG---TV 604
Qy 496 DAKT-----GIIDISNSG---KKTQALHFTSPL-----LTAGTESRERLTN 534
Db 605 NEKTYEPTQDLTAPTIVLSVTSKNGDAGLKVTEAQEFVKESENLTNFATVTVSGSITY 664
Qy 535 GKYSYINKLKFRVKNMOWTDEASSKLDPSKVI-----QVAFETEGTDEIGLIVNAKAGN 590
Db 665 QOVAVKA-----GANLSALTASDIIPASVEAVTGQDGTVKVKAAN-----706
Qy 591 DDIPVQGGKMMIDGGDGHDRVPYKDGFGFNIIVDGTSAEAGSYTVNRKVARGDYHEV 650
Db 707 -----QLERNQGYKLVVFGK-----GATAPVKDAANAN-TLATNYIYTF 745
Qy 651 VKROBTGVKKTETIYRDYELRVKGYGYQSTDNLSVEEVIGSFQFNDVPKSGKFNDIFH 710
Db 746 TEGQDVTAFTVKVKF-----GDSLKDADAV--TTLTNVDAGQKFTIQF- 787
Qy 711 SGEDDLDDGAGDRLPGKGNDRLSGDEGDDLDDGSGDDVLLGGAGNDVYIFRKGDG 770
Db 788 -----SEELKTSGS--LVGGKVT-----VEKLTTNGWVD-----AGTGTTVSVAPKTD 830
Qy 771 ND-----TLYDGTGNDKLAFADANISDIEMIERTKEGI-----IVKNDHSGSINI 817
Db 831 NGKVTAAVVTITGLDNNDK-----DAKURLVVDKSGSDGIADVAGNVIKEKD-----877
Qy 818 WYITSNLQYQS-NKTDHKEIQLIKQD-----SYITSDDQIDKILQDKDGTVITSOBL- 870
Db 878 -----ILIRYNSWRHTVASVKAADKQONASAPPTSTAD-----TTKSL 920
Qy 871 -----KKLADENKSKLASIASLNKLVGSMALFGTANVSVALQPIPTQPGILAP 925
Db 921 VEFNETDLA-EVKPENIVVKAAGN-----AVAGTVTALDGSSTNKFVFTPSQELKAG 971
Qy 926 SV 927
Db 972 TV 973
```

RESULT 39

```
US-09-882-227-522
; Sequence 522, Application US/09882227
; Publication No.: US20030158396A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20030158396A1el Helicobacter Polypeptides in the
; FILE REFERENCE: 06132/047002
; CURRENT APPLICATION NUMBER: US/09/882, 227
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/902, 615
; PRIOR FILING DATE: 1997-07-29
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 522
; LENGTH: 2893
; TYPE: PRP
; ORGANISM: Helicobacter pylori
; US-09-882-227-522
```

Query Match 3.8%; Score 175; DB 12; Length 2893;

Best Local Similarity 18.2%; Pred. No. 0.0056;

Matches 199; Conservative 149; Mismatches 378; Indels 366; Gaps 48;

```
Qy 1 MSNINVIKSN-----QAGLNS--TKSGLKNLYLAIPKDYDPQKGGTLNDFIKADE 50
Db 1843 LNQANIVSSQTDGIFSMGQEGINKVFNQAGLANILGEVAVQSKAGGLNVLNLTGS 1902
Qy 51 ---LGIARLAEEPNT-----ETAKSVDTVNQFLSLTQTGI---84
Db 1903 NSVIGGYLTPEQKNQTLSQLGQNNFDNLMNDSGLNTAIK--DLIROKLG-F-MTGLVGG 1959
Qy 85 ---AISATKLEKFLQKHSNKLAKGLDSVENIDRKLKASNVLSTLSFLGTALAGIEL 140
Db 1960 AGLGIDLQNPCKLIGSMSINDLLS-----KGLFNQITGFISANDIGOVISVWL 2009
Qy 141 DSLIKKGDAAPDALAKASIDLINEIIG-----NLSQSTQTTEAFSSOLAK--LGSTI 190
Db 2010 QDIVKPSNALKNDVAALGKQWIGFELGQDTLNSLESLLQNOQIKSVLDKVLAAKGLGPIY 2069
Qy 191 SQAKG--FSNIGNK-----LQNLNPSKTNLG-----214
Db 2070 EQGLGDLIPNLGKGLFPAPYGLSQVMQKGFSPNAQNVFQNSTFSNANGGTLNFMAGN 2129
Qy 215 -----LEIITCLLSGISAGFALADKNAKSTGKVAAGFELSNOVIGNV 256
Db 2130 SLIFAGNHHIAFTNHAGTLLQSLDQVSNIN-----ITTLNANGLKINNAANNVSVSQNL 2185
Qy 257 TKAISSVYLAQRVAAGLSTTGCAVAALITSS-----IMLAISPLAFMNAADKFN- 304
Db 2186 FVSASCAQSDPTTANFPCALSQAQSTNGCASSNNASNNAPIALSNNDESILMVAANDENF 2245
Qy 305 ---HANALDEPAK-----QPRKFGYDGDHLLAEYQ-----331
Db 2246 SGNIYANGVVDVFSKIKGSANIKNLYLNNAQFOANNLTISNOAVLEKNASFTVNNLNIQ 2305
Qy 332 ---RGVGTIEASLTITSTALGAVSAGVSAAGVSAVG-----TPIALLVA- 373
Db 2306 AFNNNATQKIEVLQNLVIASNASLSTGIYGLEVGAGLANSGA IHFNLENTQTPTPLIOAE 2365
Qy 374 GVTGLIS-----GILEASKQAMFESVANRLQKILEW-----405
Db 2366 GIINLNTQTPPMNVNNSMANNITYTLKSSRYIDYNNIPNSLOSQYLNLYTLININGHI 2425
Qy 406 EKQNGQNYF-----DKGY-----DSRYAAYLANNLKLFISELNKELE-----442
Db 2426 ECKNGALTYLQQRVLLQDKGLLSVALPNSNNAS--QNNILSLSVLYNQVMKSCGDKAMD 2483
```

```
QY 443 -----AERVIAITQORWNNIGELAG-----ITKLERIKSGKAYADAFEDGKKVEAGS 491
Db 2484 FPPPTLQDIYVIGIQQSALNQLEAVGGNAIKWLSLMMETKENPFAFIYKNSHLEIL 2543
QY 492 NITLDAKTIIDISNSN--GKKTQALHFTSPLLTAGTSRRLRNGKYSYINKLKFGRVK 549
Db 2544 GWTKDLQNTASLISNPNFRDNATNLE-----LASYTQOTSRLT-----KLS 2585
QY 550 NMQVTDGASSKLDPSKVIQVRAETEGTDEIGLIVNAKAGNDIDIPVGGKKN----- 601
Db 2586 DFRSREGS-----DPS-----LLELKNRFSDP-----NPEVFYKSQLSKHPNNLW 2629
QY 602 -----IDGGDGHDRVYFSKGGGFG-----NITVDGTSATEAGSYTVNRKVARGDIYHE 649
Db 2630 QGVGASFISSGNG-----TLYGLNAGYDRLVKNVILGGYVAYGYSDF-----NGNINHS 2679
QY 650 V-----VKR-----QETKVGKRT-----ETIQYRDYELRKVG 676
Db 2680 LGNNVDVGMARAFKRNFEFLSANETYGNATSINSSNLSLVLNQRYNYNTWTTSVNG 2739
QY 677 -YGYOSTNLKSV--EIVIGSOFNDV-FKSGKFPD-----IFHSGEADD---LLDGGAG 723
Db 2740 NYGYDFMFKQSKSVLKPQVGLSYHFGLSGMKGNDAAYKQFLMHSNPNSESVLTLNMGLE 2799
QY 724 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG-TGND- 781
Db 2800 SRKYFGKNSYFVTARLGRDRLLIKSGNTVR-FVGENTLLYRKGEVNTFASVITGGEM 2858
QY 782 ---KLAFADANI 790
Db 2859 HLWRLVYVNAV 2870

RESULT 40
US-10-193-764-63
; Sequence 63, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loomsore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT PROTEINS
; FILE REFERENCE: 1038-1239MIS
; CURRENT APPLICATION NUMBER: US/10/193,764
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 1536
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-193-764-63
```

```
Query Match 3.8%; Score 174.5; DB 12; Length 1536;
Best Local Similarity 20.8%; Pred. No. 0.0023;
Matches 138; Conservative 80; Mismatches 243; Indels 203; Gaps 28;

QY 370 LLVAGVTGLISGILEASKQAM-----FESVANRLQKILEWEKONGQNYFD 416
Db 54 LLSLGVTSIPQSVLASGLQGMVDVVGATMQVDGNKTIIRNSVDDIINWKQFNIDQN--- 110
QY 417 KGYDSRYAAYLANNIKFLSELNKELEAERVIAITQORWNNNIGELAGITKLGRKSGKA 476
Db 111 -----EMVQFLQENNSAVFNKRVTS-----NQISQLKGILD-----SNGQV 146
QY 477 Y---ADAFEDGKKVEAGSNITLDATGIIDISNSNGK-KTQALHFTSPLLTAGTSRRL 532
Db 147 FLINPNGITIGKDAINTN-GFTAST--LDISNENIKARNFTFEQTKOKALAEIVNHGLI 203
QY 533 TNGKYSYINKLKFGRVKQNV--TDGEASSKLDPSKV-IQVRAETEGTDEIGLIVNAKAG 589
```

```
Db 204 TVCKDGSVN-LIGKVKNEGIVSNGGSIILLAQKQITISDIINPTITYSIAAPENEAVN 262
QY 590 NDDIFVQGGKMNIDGGCHDRVYFSKDDGGFNITVDGTSATEAGSYTVNRKVARGDIYHE 649
Db 263 LGDIFAKGGINVRAAATIRNQ-----GKLSADSVSKDKSGNIVLSAKEGBAEIGGV 313
QY 650 V-VKRQSTKVGKRTETIQYRDYELRKVGYGVOSTDNLKSVEEVIGSOFNDVFKGSKFNDI 708
Db 314 ISAQNOAKGKGLMIT-----GDKVTLTKGAVIDLSGKEGGETYLG----- 354
QY 709 FHSGEGLDLDGGAGDRLFGGKG-----NDRLSGDE-----GD-DLLD 746
Db 355 -----GDERGEGKGIQAKKTSLEKSGTINVSKEKGGRAIVWSDIALID 400
QY 747 G-----GSGDDVLNGG-----AGNDVYI-----FRKGGOND 772
Db 401 GNINAQSGDIAKTGGFVETSGHDLFIKDNAIVDAKEWLLDPPDNVSNINAEAGRSNTSED 460
QY 773 TLYDGTGN-----DKLAFADANISDIMIERKEGIIIVKRNH-SGSINIPRWIT- 821
Db 461 DEYTGSGNSASTPKRNKEKTTLNTTLESILKKGTFVNITANQRIYVNSSINLSNGSLTL 520
QY 822 -----SNLQNTQSNKTD-HKIEQLIGKDGSVITSDQIDKIL 856
Db 521 WSEGRSGGVVEINNDITTGDDTRCANLTIYSGGWVDVHKNISLGAQGNINITAKQ--DIA 578
QY 857 QDKKDGTVITTSQELKKLADEN--KSQKLSASDIASSL-----NKLVGSMALF 901
Db 579 FEKGSNOVITQCGGTITSGNQKGFNNVSLNGTSGGLQFTTKRTNKYAITNKPFGTLNIS 638
QY 902 GTAN 905
Db 639 GKNV 642
```

Search completed: February 17, 2004, 10:18:33
Job time : 55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - protein search, using sw model

Run on: February 17, 2004, 10:09:49 ; Search time 46 Seconds

(without alignments)
3198.686 Million cell updates/sec

Title: US-10-069-799-5

Perfect score: 4647

Sequence: 1 MSNINWIKNIQGLNSTKS.....SSNALQPIPTQIGILAPSV 927

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 4647 | 100.0 | 927 | 22 | M. bovis Dalton 2d |
| 2 | 2334 | 50.2 | 1098 | 18 | Chimeric protein # |
| 3 | 2334 | 50.2 | 1098 | 21 | Bovine IL-2/Pasteu |
| 4 | 2332 | 50.2 | 1098 | 13 | Bovine IL-2 - LKT |
| 5 | 2329 | 50.1 | 953 | 14 | Leukotoxin protein |
| 6 | 2327 | 50.1 | 1098 | 15 | Bovine IL-2/LKT ch |
| 7 | 2326.5 | 50.1 | 924 | 14 | Recombinant leukot |
| 8 | 2326.5 | 50.1 | 924 | 14 | Recombinant leukot |
| 9 | 2326.5 | 50.1 | 924 | 14 | Recombinant leukot |

| | | | | | |
|----|--------|------|------|----|----------|
| 10 | 2326.5 | 50.1 | 926 | 12 | AA14482 |
| 11 | 2326.5 | 50.1 | 926 | 14 | AA134545 |
| 12 | 2326.5 | 50.1 | 926 | 15 | AA150291 |
| 13 | 2326.5 | 50.1 | 926 | 19 | AA179568 |
| 14 | 2326.5 | 50.1 | 977 | 17 | AA179568 |
| 15 | 2326.5 | 50.1 | 977 | 19 | AA179568 |
| 16 | 2326.5 | 50.1 | 1069 | 18 | AA152748 |
| 17 | 2326.5 | 50.1 | 1069 | 15 | AA113867 |
| 18 | 2326.5 | 50.1 | 1069 | 21 | AA121074 |
| 19 | 2325 | 50.0 | 953 | 11 | AA107167 |
| 20 | 2325 | 50.0 | 953 | 15 | AA160072 |
| 21 | 2325 | 50.0 | 953 | 22 | AA104638 |
| 22 | 2323.5 | 50.0 | 943 | 14 | AA134546 |
| 23 | 2318.5 | 49.9 | 936 | 14 | AA134547 |
| 24 | 2317.5 | 49.9 | 924 | 12 | AA110889 |
| 25 | 2317.5 | 49.9 | 926 | 17 | AA103945 |
| 26 | 2313 | 49.8 | 953 | 12 | AA115159 |
| 27 | 2312.5 | 49.8 | 951 | 14 | AA134548 |
| 28 | 2297.5 | 49.4 | 934 | 17 | AA107637 |
| 29 | 2236.5 | 48.1 | 956 | 12 | AA112561 |
| 30 | 2236.5 | 48.1 | 956 | 18 | AA122156 |
| 31 | 2236.5 | 48.1 | 956 | 21 | AA151410 |
| 32 | 1924 | 41.4 | 1049 | 18 | AA122159 |
| 33 | 1924 | 41.4 | 1049 | 21 | AA151412 |
| 34 | 1924 | 41.4 | 1244 | 15 | AA154781 |
| 35 | 1763.5 | 37.9 | 1022 | 18 | AA122152 |
| 36 | 1763.5 | 37.9 | 1022 | 21 | AA151406 |
| 37 | 1727 | 37.2 | 1023 | 16 | AA176991 |
| 38 | 1464.5 | 31.5 | 608 | 22 | AA104636 |
| 39 | 1376.5 | 29.6 | 758 | 17 | AA186998 |
| 40 | 1159.5 | 25.0 | 450 | 22 | AA104637 |
| 41 | 904.5 | 19.5 | 544 | 19 | AA179570 |
| 42 | 901 | 19.4 | 695 | 19 | AA179573 |
| 43 | 901 | 19.4 | 695 | 21 | AA158361 |
| 44 | 901 | 19.4 | 695 | 21 | AA158133 |
| 45 | 897.5 | 19.3 | 490 | 20 | AA133929 |

ALIGNMENTS

RESULT 1

AA162110

ID AA162110 standard; Protein; 927 AA.

XX AA162110;

AC AA162110;

XX 29-MAY-2001 (first entry)

DT 29-MAY-2001 (first entry)

DE M. bovis Dalton 2d RTX toxin A subunit.

XX Moraxella; antigen; immune response; infection; RTX toxin; vaccine;

KW antibacterial; A subunit.

XX Moraxella bovis.

OS Moraxella bovis.

XX WO200116172-A1.

FN WO200116172-A1.

XX 08-MAR-2001.

XX 31-AUG-2000; 2000WO-AU01048.

XX 31-AUG-1999; 99AU-0002571.

XX (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX (UYME) UNIV MELBOURNE.

XX Farn J, Strugnell R, Tennent J;

XX WPI; 2001-235092/24.

XX N-PSDB; AAF57290.

XX Novel Moraxella bovis antigen useful in compositions for raising immune

Leukotoxin 352 pro

Recombinant leukot

Leukotoxin 352 pol

LKT-GnRH protein f

LKT-GnRH chimeric

Bovine IFNgamma/LK

Chimeric protein #

Bovine gamma-IFN/p

10kD PTX protein

PTX protein of Pa

Pasteurella haemol

Somatostatin-leuko

GnRH-leukotoxin ge

Leukotoxin 352 enc

P. haemolytica tru

Leukotoxin from P.

Rotavirus VP4-leuk

P. suis leukotoxin

APPA haemolysin an

ApXII protein. A

A. pleuropneumonia

ApXIIIB protein.

A. pleuropneumonia

Leukotoxin ApplIIA

ApXIA protein. Ac

A. pleuropneumonia

LhaA (low homology

Pasteurella haemol

Enterohaemorrhagic

Pasteurella haemol

LKT-GnRH chimeric

LKT-GnRH chimeric

Leukotoxin/gonadot

Gonadotropin relea

Leukotoxin carrier

PT response in an animal, has protease, lipase or hemolysin activity -

XX PS Claim 26; Fig 5; 60pp; English.

XX CC The invention relates to new Moraxella bovis antigens and nucleic acid

CC sequences encoding these antigenic polypeptides. The antigenic

CC polypeptides and polynucleotides are useful for raising an immune

CC response in an animal directed against Moraxella, preferably against

CC M. bovis or M. catarrhalis, and for treating Moraxella infections. The

CC present sequence represents the amino acid sequence of the A subunit of

CC the RTX toxin from M. bovis Dalton 2d.

XX Sequence 927 AA;

SQ Query Match 100.0%; Score 4647; DB 22; Length 927;

Best Local Similarity 100.0%; Pred. No. 4.1e-281;

Matches 927; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSNINVIKSNIOAGLNSTKSGLNLYLAIPKDYDPQKGGTLDNFKADELGIAARLAE 60

DB 1 MSNINVIKSNIOAGLNSTKSGLNLYLAIPKDYDPQKGGTLDNFKADELGIAARLAE 60

QY 61 NHTETAKSVDTVNOFLSLTGTGIAISATKLEKFLQKHSNKLAKGLDSVENIDRKLGA 120

DB 61 NHTETAKSVDTVNOFLSLTGTGIAISATKLEKFLQKHSNKLAKGLDSVENIDRKLGA 120

QY 121 SNVLSTSSPFTALAGIELSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAF 180

DB 121 SNVLSTSSPFTALAGIELSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAF 180

QY 181 SOLAKGSTISQAKGFSNIGNKLNQNLNFSKTNLGLIITGLSGISAGFALADKNASTGK 240

DB 181 SOLAKGSTISQAKGFSNIGNKLNQNLNFSKTNLGLIITGLSGISAGFALADKNASTGK 240

QY 241 KVAAGFELSNQVIGNVTKAISYLAQVAAGLSTTGAVALITSSIMLAISPLAFWNA 300

DB 241 KVAAGFELSNQVIGNVTKAISYLAQVAAGLSTTGAVALITSSIMLAISPLAFWNA 300

QY 301 DKFNHANALDFAKQFRKFGYDGHLLAAYQGVGTTEASLTITSTALGAVSAGVAAA 360

DB 301 DKFNHANALDFAKQFRKFGYDGHLLAAYQGVGTTEASLTITSTALGAVSAGVAAA 360

QY 361 GSAVGTPIALLVAGVTGLISGLEASKQAFESVANRLQGLKEWKGONQYDFDGYD 420

DB 361 GSAVGTPIALLVAGVTGLISGLEASKQAFESVANRLQGLKEWKGONQYDFDGYD 420

QY 421 SRYAAYLANNLKFLSELNKELEAERVIAITQORWNNIGELAGITKLERIKSGKAYADA 480

DB 421 SRYAAYLANNLKFLSELNKELEAERVIAITQORWNNIGELAGITKLERIKSGKAYADA 480

QY 481 PEDGKVEAGSNITLDAGTIIIDISNSNGKKTQALHFTSPLLTAGTSRRLTNGKYSYI 540

DB 481 PEDGKVEAGSNITLDAGTIIIDISNSNGKKTQALHFTSPLLTAGTSRRLTNGKYSYI 540

QY 541 NKLKFRVKNQVDTGEASSKLDPSKIVQVAETEGTDEIGLIYNKAGNDDIFVGOKM 600

DB 541 NKLKFRVKNQVDTGEASSKLDPSKIVQVAETEGTDEIGLIYNKAGNDDIFVGOKM 600

QY 601 NIDGGDGHDRVYFYSKDGFGNITVDGTSATBAGSVTVNRKVARGDIYHEVVKRQETKVGK 660

DB 601 NIDGGDGHDRVYFYSKDGFGNITVDGTSATBAGSVTVNRKVARGDIYHEVVKRQETKVGK 660

QY 661 RTETIQYRDYELRKVGYQSTDLNLSVEEVIGSQFNDVFKGSFNDIFHSGEGDLDLDG 720

DB 661 RTETIQYRDYELRKVGYQSTDLNLSVEEVIGSQFNDVFKGSFNDIFHSGEGDLDLDG 720

QY 721 GAGDRFLFGKGNDRSLDGEDDLDLGGSGDDVLNGGAGNEVYIFRKGDDGNDTLYDGTGN 780

DB 721 GAGDRFLFGKGNDRSLDGEDDLDLGGSGDDVLNGGAGNEVYIFRKGDDGNDTLYDGTGN 780

QY 781 DKLAFADANISDIIMERTKEGIIIVKRNHDSGSINIPRWYIISNLQYOSNKTDKHIEQLI 840

DB 781 DKLAFADANISDIIMERTKEGIIIVKRNHDSGSINIPRWYIISNLQYOSNKTDKHIEQLI 840

QY 841 GKDGSYITSDIQDKILODKKQGTVTITSQELKKLADENKSKLSASDIASSINKLVGSMAL 900

DB 841 GKDGSYITSDIQDKILODKKQGTVTITSQELKKLADENKSKLSASDIASSINKLVGSMAL 900

QY 901 FGTVANSVSSNALQPIPTOPTOGILAPSV 927

DB 901 FGTVANSVSSNALQPIPTOPTOGILAPSV 927

RESULT 2

AAW13866
ID AAW13866 standard; Protein; 1098 AA.

XX AAW13866;

AC AC

DT 25-MAR-2003 (updated)

DT 12-MAY-1997 (first entry)

XX Chimeric protein #1.

XX RTX cytotoxin; cytokine; immunogen; chimeric protein; cytokine; vaccine;
interleukin-2; IL-2; gamma interferon; gamma IFN; leukotoxin; pneumonia;
Pasteurella haemolytica; LKT352; respiratory disease; shipping fever;
fibrinous pneumonia; cattle; therapy.

OS Synthetic.

XX US5594107-A.

XX 14-JAN-1997.

XX 20-DEC-1993; 93US-0170126.

XX 20-DEC-1993; 93US-0170126.

XX 22-AUG-1990; 90US-0571301.

XX 16-OCT-1991; 91US-0777115.

XX (CIBA) CIBA GEIGY CANADA LTD.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Campos M, Hughes HPA, Potter A;

XX WPI; 1997-099529/09.

XX N-PSDB; AAT60032.

XX Immunogenic chimeric proteins comprising cytokine linked to RTX
toxin - useful in vaccines, esp. against shipping fever in cattle

XX Claim 10; Column 25-32; 56pp; English.

XX AAW13866 and AAW13867 represent immunogenic chimeric proteins of the
invention. This sequence represents a chimeric protein containing the
bovine interleukin-2 (IL-2) sequence and a leukotoxin sequence. The
chimeric proteins of the invention comprise a cytokine, selected from
IL-2 and gamma interferon (gamma IFN), linked to at least one RTX toxin
epitope (preferably the sequence shown in AAW13865). The RTX toxin used
to provide the epitope sequence is preferably a leukotoxin, especially
the full-length Pasteurella haemolytica leukotoxin. Alternatively, the
leukotoxin is a truncated leukotoxin lacking leukotoxic activity,
especially LKT352. The chimeric proteins can be used for the production
of vaccines against respiratory diseases such as pneumonia, particularly
fibrinous pneumonia caused by P.haemolytica, including shipping fever in
cattle.

XX (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 1098 AA;

Query Match 50.2%; Score 2334; DB 18; Length 1098;

Best Local Similarity 50.1%; Pred. No. 7.4e-137;

Matches 465; Conservative 175; Mismatches 263; Indels 26; Gaps 13;

QY 8 KSNIOAGLNSTKSGLNLYLAIPKD--YDPQKGGTLDNFKADELGIAARLAEPPNHTET 65

Db 173 QSLTQAG-SLKTGAKKIILYIPQNYDYDEQNGLDLVKAAEELGIEVQREERNIAT 231
Qy 66 AKKSVPTVNOFLSLTQGTGIAISATKLEFLQKHSNTKLAAGLSDVENIDRKLKASNVLS 125
Db 232 AQTSLGTIQTALGIELSLIKKGDAPDALAKASIDLINEIIGLSQSTQTIASFSSQLAK 288
Qy 126 TLSSFLGTALAGIELSLIKKGDAPDALAKASIDLINEIIGLSQSTQTIASFSSQLAK 185
Db 289 GIOSILGSLVAGMDLDEAL-QNNSNHQALAKAGLELTNSLIENIANSVKTLDFFGQISQ 347
Qy 186 LGSTISQAKGFSNIGNKQLNL-NFSKTNLGLLEITGLLSGISAGFALADKNASTGKVVAA 244
Db 348 FGSKLQNIKGLTGLDKLKNIGGLDKAGLGLDIVISGLLSGATAALVLDKNASTAKKVA 407
Qy 245 GFELSNQVIGNTKATISSYVLAQRVAAGLSTGTAVALITSSIMLAISPLAFNAADKFN 304
Db 408 GFELANQVGNITKAVSSYTLAQRVAAGLSTGTPVAALIASVLSLAISPLAFAGIADKEN 467
Qy 305 HANALDEFKQFRKFGYDGHLLAEYQGVGTIEASTTTISTALGAVSAGVSAAGVSAV 364
Db 468 HAKSLESYAEFRKGLYDGNLNLAEYQGVGTIDASVTAINTALAAIAGGVSAAGVSAV 527
Qy 365 GTPIALVAGVTGLISGILEASQAMPESVANRLQKILEWEKONGGQNYFDKGYDSRYA 424
Db 528 ASPIALLVSGITGVISTILQYSQKMPFHVANKIHNIKEVKNHKGKYNFENGYDARYL 587
Qy 425 AYLANNLKFLSELNKELEAEVIAITQORWNNIGELAGITKGERIKGKAYADAPEDG 484
Db 588 ANLQNMKFLNLNKELEAEVIAITQOQWNNIGDLGAGISRLGKVKLSGKAVDAFEFG 647
Qy 485 KVEAGSNITLDKATGIIDNSNGKKTQALHTFSPLLTAGTSREBLTNGKYSYINKL 544
Db 648 KHAKADKLVLQDSANGIIDVSNKSKAKTOHLFRTPLLTGTEHREVRQVKGVEYITKLN 707
Qy 545 FGRVKNQVTDGEASSKLDPSKVIQV-----AETEGTDEIGLI VNAKAGNDDIPVQ 597
Db 708 INRVDSNKTGDAASSTFDLTNVVQRIEGLDNAGNTKTKETKI IAKLGGEDNVFVGS 767
Qy 598 GQWIDGGDGHDRVYFSKDGFGNITVDGTSATEAGSYTVNRKVAFGDIHEVYKQETK 657
Db 768 GTTEIDGEGYDRVHYSR-GNYGALTIDATKETEGSYTVNRVETGKALHEVTSHTAL 826
Qy 658 VGRKTEIQRDYELRVKGVGYQSTDMKSVVEVIGSQFNDVPKGSKFNDI FHSBGDDL 717
Db 827 VGNREKIEYR-HSNNHQAHGYTKTLKAVEEIIGTSHNDIPKGSFNDAPFNGDGVDT 885
Qy 718 LDGAGADRLFGKGNDRSLGDEGDLGSGGDDVLNGAGNDVYIFRKGDDGNDLTLDG 777
Db 886 IDGNDGNDRLFGKGGDDILGNGGDDFDGKGNDLHGGKDDIFVRKGGNDIITDS 945
Qy 778 TGNDKLAPADANTISDIMERTEGIIIVKRDHSGSINIPRWY-----ITSNLQYQSNKTD 833
Db 946 DGNDKLSFSDSNLKDITFEKVKNLVI-TNSKKEKVTIQNWFEADPAKEVPNYKATK-D 1003
Qy 834 HKIEQLTGKGSYITSDIKILQDKGCVITFSQELKADENKSKLASDIASSLNK 893
Db 1004 EKIEEIIQNGERTISQVDDLI--AKNGKXITODELSKVVNDVYELLKHS-KNVTNSLDK 1060
Qy 894 LVGSMALFGTANSVSSNALQPIPTQGI 922
Db 1061 LISSVAFSTSSNDRNVIVAPTSMLDQSL 1089

RESULT 3

AAB21073

ID AAB21073 standard; Protein; 1098 AA.

XX AC AAB21073;

XX DT 19-DEC-2000 (first entry)

XX DE Bovine IL-2/Pasteurella haemolytica leukotoxin fusion protein.

XX KW Bovine IL-2; interleukin-2; leukotoxin; LKT; respiratory disease; pneumonia; shipping fever; cattle; livestock; anti-Pasteurella vaccine; immunogen.
XX OS Chimeric - Bos taurus.
XX OS Chimeric - Pasteurella haemolytica.
XX PA (UUSA-) UNIV SASKATCHEWAN.
XX PA (CIBA) CIBA GEIGY CANADA LTD.
XX PI Campos M, Hughes HPA, Potter A;
XX DR WPI: 2000-531543/48.
XX DR N-PSDB; AAA72483.
XX PT Vaccine for stimulating immunity against pneumonia comprises chimeric protein comprising gamma-interferon and leukotoxin derived from Pasteurella haemolytica
XX PS Example 1; Column 31-38; 56pp; English.
XX CC The invention relates to a novel vaccine composition comprising an immunogenic chimeric protein that comprises gamma-interferon (gamma-IFN) or an active fragment thereof, linked to an epitope of a Pasteurella haemolytica leukotoxin (LKT). Pasteurella species, especially Pasteurella haemolytica, are responsible for respiratory diseases in a range of agricultural animals, most particularly cattle, but also sheep, pigs, horses and fowl. Shipping fever is the most economically important respiratory disease associated with Pasteurella species, affecting 15-30% of exposed cattle and resulting in a 2-5% mortality rate in the exposed population. The vaccine composition of the invention is useful for preventing or ameliorating respiratory diseases such as pneumonia, particularly shipping fever pneumonia, in livestock. The present sequence represents a fusion protein comprising bovine interleukin-2 (IL-2) and Pasteurella haemolytica leukotoxin, which may also be used as an anti-Pasteurella vaccine.
XX SQ Sequence 1098 AA;
Query Match 50.2%; Score 2334; DB 21; Length 1098;
Best Local Similarity 50.1%; Pred. No. 7.4e-137;
Matches 465; Conservative 175; Mismatches 263; Indels 26; Gaps 13;
Qy 8 KSNIQAGLSTKSLKXNLVLAIPKD--YDPKGGTLDNDFIKADELGIARLAEPNHET 65
Db 173 QSLTQAG-SLKTGAKKIILYIPQNYDYDEQNGLDLVKAAEELGIEVQREERNIAT 231
Qy 66 AKKSVPTVNOFLSLTQGTGIAISATKLEFLQKHSNTKLAAGLSDVENIDRKLKASNVLS 125
Db 232 AQTSLGTIQTALGIELSLIKKGDAPDALAKASIDLINEIIGLSQSTQTIASFSSQLAK 288
Qy 126 TLSSFLGTALAGIELSLIKKGDAPDALAKASIDLINEIIGLSQSTQTIASFSSQLAK 185
Db 289 GIOSILGSLVAGMDLDEAL-QNNSNHQALAKAGLELTNSLIENIANSVKTLDFFGQISQ 347
Qy 186 LGSTISQAKGFSNIGNKQLNL-NFSKTNLGLLEITGLLSGISAGFALADKNASTGKVVAA 244
Db 348 FGSKLQNIKGLTGLDKLKNIGGLDKAGLGLDIVISGLLSGATAALVLDKNASTAKKVA 407
Qy 245 GFELSNQVIGNTKATISSYVLAQRVAAGLSTGTAVALITSSIMLAISPLAFNAADKFN 304

Db 408 GFELANQVGNITKAVSSYILQORVAAGLSSTGPVAAALIASTVSLAISPLAFAGIADKFN 467
 Qy 305 HANALDEFKQFRFGYDGDHLLAEYQGVCTTIEASLTITSTALGAVSAGVSAAGVSAV 364
 Db 468 HAKSLESYAEFRFKGLYDGDNLAEYQGVCTTIDASVTAINALAAAGVSAAGVSAV 527
 Qy 365 GTPIALVAVGTGLISGLEASKAMPESVANRQGLKLEWKGKNGGQNYFDKGYDSRYA 424
 Db 528 ASPIALVSGITGVISTILQYKQAMPEHVAANKHKNKIVEMKNNHGNFENGVDARYL 587
 Qy 425 AYLANNLKFLSELNKELEAEERVIATTOQRWNNNIGELAGITKLGRIKSGKAYADAFEDG 484
 Db 588 ANLQDNMKFLNLNKLQAEERVIATTOQRWNNNIGELAGISRLGEKVLGSKAYADAFEEG 647
 Qy 485 KKVAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRRLTGKYSYINKLK 544
 Db 648 KHKADKLQVLDANGIIVDSNSGKAKTQHILFRTPLLTGTEHRRVQTKGYEYITKLN 707
 Qy 545 FGRVKNQVTDGEASSKLDPSKVIRV-----AETEGTDEIGLIVNAKAGNDDIFVQ 597
 Db 708 INRVDSWKITDGAASSTFDLTNNVQIRIGIELDNAGNVTKTKETKIIAKLGEAGDNNVFGS 767
 Qy 598 GKMNIDGGDHRDVFYSKGGFNGITVDGTSATBAGSYTVNRKVARGDIYHEVVKROETK 657
 Db 768 GTTEIDGEGYDRVHYSR-GNYGALTIDATKETEGSYTVNRFVETGKALHEVTSHTAL 826
 Qy 658 VGKTTETQYRDYELRVKGYQSTDNLSKVEEIVGSOQFNDVFKGSKFENDIFHSGEGDDL 717
 Db 827 VGNREEKLEYR-HSNQNHAGYTKDTLKAVEEIIGTSHNDIFKSGKFNDAFNGGVD 885
 Qy 718 LDGGAGDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLLGGAGNDVIFRKGDNNTLYDG 777
 Db 886 IDGNDGNRLFGGKGNDRLSGDEGDDLLDGGSGDDVLLGGAGNDVIFRKGDNNTLYDG 945
 Qy 778 TGNDKLAFADANISDIEMIERTKEGIIVKRNDSHSGSINIPRWY----ITSNLQNYQSNKTD 833
 Db 946 DGNDKLSPSDSNLKDLTPEKVKHNLVI-TNSKKEKVTIQNFREADFAKEPVNKATK-D 1003
 Qy 834 HKIQOLIGKDSYITSDQIDKILQDKDGTVITSGELKKLADENKSOKLSASDIASSLNK 893
 Db 1004 EKIEEIIQNGERITSKQVDDLI--AKNGKITQDELKVVVDNYELLKHS-KNVTNSLDK 1060
 Qy 894 LVGSMALFGTANSVSSNALOPTQGTQGI 922
 Db 1061 LISSVSAFTSSNDRNLVAPTSMLDQSL 1089

RESULT 4

AAR22103

ID AAR22103 standard; Protein; 1098 AA.

XX AC AAR22103;

XX DT 06-JUL-1992 (first entry)

XX DE Bovine IL-2 - LKT fusion protein.

XX KW Interleukin 2; leuko-toxin; vaccine; pneumonia; respiratory

XX OS Pasteurella haemolytica.

XX OS Bos taurus.

XX FN W09203558-A.

XX PD 05-MAR-1992.

XX PF 22-AUG-1991; 91WO-CA00299.

XX PR 22-AUG-1990; 90US-0571301.

XX PA (POTT/) POTTER A.

XX XX

PI Potter A, Campos M, Hughes HPA;

XX WPI; 1992-096901/12.

DR N-PSDB; AAQ22771.

XX Interleukin 2-leuko-toxin gene fusion - encodes fusion protein

PT useful as vaccine for animal pneumonia

XX Claim 20; Fig 3; 68pp; English.

PS The IL-2-LKT protein was encoded by a chimeric gene contg. the

XX bovine IL-2 gene fused to the DNA encoding at least one epitope of

CC leukotoxin from P. haemolytica. IL-2-LKT was gel purified and

CC ligated into the expression vector pGH433 lacI. The resulting

CC clone pAA356 (APCC 68986) contd. the desired gene fusion under the

CC control of the E. coli lac promoter. The protein produced by the

CC gene fusion is useful in a vaccine compen. with a pharmaceutically

CC acceptable vehicle, e.g. a carrier homologous to a rotavirus VP6

CC inner capsid protein. The vaccine can be used for preventing or

CC ameliorating respiratory diseases in animals e.g. shipping fever

CC or pneumonia.

CC See also AAR24124,5.

XX Sequence 1098 AA;

SQ Query Match 50.2%; Score 2332; DB 13; Length 1098;

Best Local Similarity 50.1%; Pred. No. 9.9e-137; Indels 26; Gaps 13;

Matches 465; Conservative 174; Mismatches 264;

Qy 8 KSNITAGLNSTKGLNKLAIAPKD--YDPQKGTGLNDFIKAADELGIARLAEPNHET 65

Db 173 QSLTQAG-SSLTKGAKKIIILYIQNYQYDEQNGQLDLVKAEEELGIEVREERNIAT 231

Qy 66 AKSVDTVNOFLSITQTGIAISATKLEKFLQKISTNKLAKGLSDVENIDRLKGVSNLS 125

Db 232 AOTSLSGTIOTAIUTERGIVLSAPQIDKLQK--TKAGQALGSAESIVONANKAKTVLS 288

Qy 126 TLFSGTGTALAGIELDSLKKGDAAPDALAKASIDILNIEIIGLSQSTOTIEAFSSOLAK 185

Db 289 GIGSILGSVUAGMDLDEAL--QNNNSQHAKAGAGLENTSLIENIANSVKILDEGEQISQ 347

Qy 186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLLEITGLLSGISAGFALADKNASTGKVA 244

Db 348 FGSKLQNIKGLTGLDKLKNIGLDRAGLGLDVLISGLSGATAALVLADKNASTAKVGA 407

Qy 245 GFELSNQVIGNVTKATISSYVLAORVAAGLSITTCVAAALITSSIMLAISPLAFWAAKFN 304

Db 408 GFELANQVGNITKAVSSYILQORVAAGLSSTGPVAAALIASTVSLAISPLAFAGIADKFN 467

Qy 305 HANALDEFKQFRFGYDGDHLLAEYQGVCTTIEASLTITSTALGAVSAGVSAAGVSAV 364

Db 468 HAKSLESYAEFRFKGLYDGDNLAEYQGVCTTIDASVTAINALAAAGVSAAGVSAV 527

Qy 365 GTPIALVAVGTGLISGLEASKAMPESVANRQGLKLEWKGKNGGQNYFDKGYDSRYA 424

Db 528 ASPIALVSGITGVISTILQYKQAMPEHVAANKHKNKIVEMKNNHGNFENGVDARYL 587

Qy 425 AYLANNLKFLSELNKELEAEERVIATTOQRWNNNIGELAGITKLGRIKSGKAYADAFEDG 484

Db 588 ANLQDNMKFLNLNKLQAEERVIATTOQRWNNNIGELAGISRLGEKVLGSKAYADAFEEG 647

Qy 485 KKVAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRRLTGKYSYINKLK 544

Db 648 KHKADKLQVLDANGIIVDSNSGKAKTQHILFRTPLLTGTEHRRVQTKGYEYITKLN 707

Qy 545 FGRVKNQVTDGEASSKLDPSKVIRV-----AETEGTDEIGLIVNAKAGNDDIFVQ 597

Db 708 INRVDSWKITDGAASSTFDLTNNVQIRIGIELDNAGNVTKTKETKIIAKLGEAGDNNVFGS 767

Qy 598 GKMNIDGGDHRDVFYSKGGFNGITVDGTSATBAGSYTVNRKVARGDIYHEVVKROETK 657

Db 768 GTTEIDGEGYDRVHYSR-GNYGALTIDATKETEGSYTVNRFVETGKALHEVTSHTAL 826

658 VGRKRTIQRDYELRKVGYGOSTDNLKSVVEVIGSQFNDVFKSGKFNDIFHSBGDDL 717
 827 VGNREKIEYR-HSNQHGAGYTKDTLKAVEIIGTSHNDIFKSGKFNDVFKSGGVD 885
 718 LDGAGDRLFGGKGNDRSLGDEGDDLLDGGSGDDVLLGAGNDVYIFRKGNDTL 777
 886 IDGNDGNDRFLFGGKGNDRSLGDEGDDLLDGGSGDDVLLGAGNDVYIFRKGND 945
 778 TGNDKLAFADANISDIEMIERTKEGIIIVKRNDSHSGSINIPRWY----ITS 833
 946 DGNKLSFSDSNFQDLTFKVKHNLVI-TNSKKEKVTIQNWFEADPAKEVPNYK 1003
 834 HKTEQLIGKDGSYITSDQIDKILQDKDGTIVTSQELKLDENKSKLSASDIAS 893
 1004 EKIEEIIQNGERITSKQVDDLI--AKNGKTIQDELKSKVDNYELLKHS-K 1060
 894 LVGSMALFGTANSVSSNALQIPITQGTGI 922
 1061 LISSVSAFTSSNDRNVLVAPTSMLDQSL 1089

RESULT 5
 AAR43865
 ID AAR43865 standard; Protein; 953 AA.
 AC AAR43865;
 DT 25-MAR-2003 (updated)
 DT 22-DEC-1993 (first entry)
 XX Leukotoxin protein.
 XX Antigenic determinants; peptides; vaccine; supernatant; leukotoxin.
 XX Pasteurella haemolytica.
 CA2081950-A.
 PD 02-MAY-1993.
 XX 02-NOV-1992; 92CA-2081950.
 XX 01-NOV-1991; 91US-0786662.
 XX (UYGU-) UNIV GUELPH.
 PI Conlon JA, Lo RYC, Shewen PE, Strathdee CA;
 WPI: 1993-227822/29.
 N-PSDB; AAQ46410.
 Vaccine for preventing Pasteurella haemolytica infections -
 contains leukotoxin protein antigenic determinants and P.
 haemolytica culture supernatant
 Claim 1; Figure 1; 40pp; English.
 One or more biologically pure antigenic determinants of the
 leukotoxin protein comprising of at least six amino acids is a
 component of a new vaccine for animals. The vaccine elicits an
 enhanced immune response after challenge with Pasteurella
 haemolytica. The vaccines other component is a bacterial free
 culture supernatant derived from a culture of Pasteurella
 haemolytica.
 (Updated on 25-MAR-2003 to correct PF field.)

Sequence 953 AA;

Query Match 50.1%; Score 2329; DB 14; Length 953;
 Best Local Similarity 49.9%; Pred. No. 1.3e-136;
 Matches 464; Conservative 175; Mismatches 264; Indels 26; Gaps 13;

8 KSNIAQAGLNSTKSLGNLYLAIPKD--YDPQKGGTLNDFIKADELGIARLAEPEPHNET 65

Db 28 QSLTQAG-SSLTKGAKIILYIPQNYQYDEQNGQLDLVKAABELGIEVQREERNIAT 86
 QY 66 AKSVDTVQVFLSLTGTGIAISATKLEKFLKHSTNKLAKGLDSDVENIDRKLGKASNVLS 125
 Db 87 AOTSLSGTIQTALGTERGIVLSAPQDKLQK---TKAGQALGSAESI VQNAKAKTVLS 143
 QY 126 TLSSFGTALAGTLDLSLKKGDAAPDALAKASIDLINEIIGNLSTQSTOTIEAFSSQLAK 185
 Db 144 GTQSLIGSLVAGWDLDEAL-QNNSNOHALAKAGLELTNSLIENIANSVKTLDFEGQISQ 202
 QY 186 LGSTISQAKGFNIGNKQLNL-NFSKTNIGLBEIITGLLSGISAGPALAKDNASTGKKVAA 244
 Db 203 FGSKLQNIKGLGTGLGDKLKNIGGLDKAGLGLDVISGLSCATAALVLADKNASTAKKVG 262
 QY 245 GFELSNQVGNVTKATSSVLAQVAAAGLSTTCGAAVALITSSIMLAISPLAFNAADKEN 304
 Db 263 GFELANQVGNVTKAVSSYLAQVAAAGLSTTCGAAVALITSSIMLAISPLAFNAADKEN 322
 QY 305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTITSTALGAVSAGVSAAGVSAV 364
 Db 323 HAKLSYAEAFKGLGYDGDHLLAEYQRGVGTIEASLTITSTALGAVSAGVSAAGVSAV 382
 QY 365 GTPIALLVAGVTGLISGLEASQAAMPESVANRLOKILEWEKQNGQNYFDKGYDSRYA 424
 Db 383 ASPIALLVSGITGVISTILQYSQAAMFEHVANKHNKI VEWKKNHGNKGFENGYDARYL 442
 QY 425 AYLANNLKLSELNKELEAEVIAITQORWNNIGELAGITKLGERIKSKAYADAFEDG 484
 Db 443 ANLQDNMKFLNLNKELOAEVIAITQOQWNNIGLAGISRLGEKVLGSKAVDVPFEG 502
 QY 485 KVEAGSNTLDAKTGIIIDISNSNGKKTQALHFTSPLLTAGTRESRLTNGKTSYINKLK 544
 Db 503 KHKADKLVQDSANGIIDVNSGKAKTQHILPRTPLTPGTEHREVRVQTKYEYITKLN 562
 QY 545 FGRVKWQVTDGEASSKLDPSKVIQV-----AETEGTDEIGLIVNAKAGNDIDFVQG 597
 Db 563 INRVDSWKITDGAASSTFDLTNNVQRIEGLDNAGNVTKTKETKI IAKLGEQDNDVFGS 622
 QY 598 GKNNIDGGCHDRVFPYSKGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKQETK 657
 Db 623 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQQSYTVNRVETGKALHEVTSHTAL 681
 QY 658 VGRKRTIQRDYELRKVGYGOSTDNLKSVVEVIGSQFNDVFKSGKFNDIFHSBGDDL 717
 Db 682 VGNREKIEYR-HSNQHGAGYTKDTLKAVEIIGTSHNDIFKSGKFNDVFKSGGVD 740
 QY 718 LDGAGDRLFGGKGNDRSLGDEGDDLLDGGSGDDVLLGAGNDVYIFRKGNDTL 777
 Db 741 IDGNDGNDRFLFGGKGNDRSLGDEGDDLLDGGSGDDVLLGAGNDVYIFRKGND 800
 QY 778 TGNDKLAFADANISDIEMIERTKEGIIIVKRNDSHSGSINIPRWY----ITS 833
 Db 801 DGNKLSFSDSNLQDLTFKVKHNLVI-TNSKKEKVTIQNWFEADPAKEVPNYKATK-D 858
 QY 834 HKTEQLIGKDGSYITSDQIDKILQDKDGTIVTSQELKLDENKSKLSASDIAS 893
 Db 859 EKIEEIIQNGERITSKQVDDLI--AKNGKTIQDELKSKVDNYELLKHS-K 915
 QY 894 LVGSMALFGTANSVSSNALQIPITQGTGI 922
 Db 916 LISSVSAFTSSNDRNVLVAPTSMLDQSL 944

RESULT 6
 AAR52747
 ID AAR52747 standard; Protein; 1098 AA.
 XX AAR52747;
 AC AAR52747;
 DT 25-MAR-2003 (updated)
 DT 01-JUL-1994 (first entry)
 XX

| | | |
|-----------------------|-----|--|
| Query Match | | 50.1%; Score 2326.5; DB 14; Length 924; |
| Best Local Similarity | | 50.0%; Pred. No. 1.7e-136; Indels 25; Gaps 12; |
| Matches | | 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12; |
| Qy | 15 | LNSTKSGKLNLYLAIPKD--YDPQKGTGLNDFIKAADELGIARLAEPNHTETAKKSVDT 72 |
| Db | 7 | LSFPKTKAKKIILYIPQNYQYDTQGNGLQDLVKAAEELGIEVQREERNIATAQTSLGT 66 |
| Qy | 73 | VNQFLSLTGTGIAISATKLEFLQKHSNTKLAKGLDSVENIDRKLKASNVLSLSSFLG 132 |
| Db | 67 | IQTALGTERGIVLSAPQIDKLQK---TKAGALGSAESIVQNVANKAKTVLSGIQILG 123 |
| Qy | 133 | TALAGIELDSLKKGDAAADALAKASIDLINEIGNLSQSTQTIEAFSSQAKLGSTISQ 192 |
| Db | 124 | SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDFEGEQTISQFGSKLQ 182 |
| Qy | 193 | AKGFSNIGNKLQNL-NFSKTNLGLIEITGLSLGSAFALADKNASTGKVAAGFELSNO 251 |
| Db | 183 | IKGLGTIGDKLNIGGLDKAGLGLDVISLGSATAALVLADKNASTAKKVGAGFELANQ 242 |
| Query Match | | 50.1%; Score 2326.5; DB 14; Length 924; |
| Best Local Similarity | | 50.0%; Pred. No. 1.7e-136; Indels 25; Gaps 12; |
| Matches | | 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12; |
| Qy | 15 | LNSTKSGKLNLYLAIPKD--YDPQKGTGLNDFIKAADELGIARLAEPNHTETAKKSVDT 72 |
| Db | 7 | LSFPKTKAKKIILYIPQNYQYDTQGNGLQDLVKAAEELGIEVQREERNIATAQTSLGT 66 |
| Qy | 73 | VNQFLSLTGTGIAISATKLEFLQKHSNTKLAKGLDSVENIDRKLKASNVLSLSSFLG 132 |
| Db | 67 | IQTALGTERGIVLSAPQIDKLQK---TKAGALGSAESIVQNVANKAKTVLSGIQILG 123 |
| Qy | 133 | TALAGIELDSLKKGDAAADALAKASIDLINEIGNLSQSTQTIEAFSSQAKLGSTISQ 192 |
| Db | 124 | SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDFEGEQTISQFGSKLQ 182 |
| Qy | 193 | AKGFSNIGNKLQNL-NFSKTNLGLIEITGLSLGSAFALADKNASTGKVAAGFELSNO 251 |
| Db | 183 | IKGLGTIGDKLNIGGLDKAGLGLDVISLGSATAALVLADKNASTAKKVGAGFELANQ 242 |
| Qy | 252 | VIGNVTKAISYVLAQRVAAGLSTGVAALITSSIMLAISPLAFMAADKFNHANALDE 311 |
| Db | 243 | VVGNIITKAVSYVLAQRVAAGLSTGVAALITSSIMLAISPLAFMAADKFNHANALDE 302 |
| Qy | 312 | FAKQFRKFGYDGHLLAEYQRGVGTIEASLTITSTALGAVSAGVSAAGVAVGTPIAL 371 |
| Db | 303 | YAEKFKLGYDGNLLAEYQRGVGTITDASVTAINALAAIAGGVSAAGVAVGTPIAL 362 |
| Qy | 372 | VAGVTGLISGLEASKQAMFESVANRLOQKILEWKQNGQNYDFKGYDSRYAAYLANNL 431 |
| Db | 363 | VSGITGVISTILQYKQAMFESVANRLOQKILEWKQNGQNYDFKGYDSRYAAYLANNL 422 |
| Qy | 432 | KPLSELNKLBAERVAITQORWONNIGELAGITKLGRIKSGKAYADAPEDGKKVAGS 491 |
| Db | 423 | KFLNLNKLBAERVAITQORWONNIGELAGITKLGRIKSGKAYADAPEDGKKVAGS 482 |
| Qy | 492 | NITLDAKGTIIDNSNGKKTQALHFTSPILTAGTESRERLTNGKSYINKLKEGRVKNW 551 |
| Db | 483 | LVQLDSANGIIDNSNGKKTQALHFTSPILTAGTESRERLTNGKSYINKLKEGRVKNW 542 |
| Qy | 552 | QVTDGEASKLDFSKVIQV-----AETGCTDEIGLIVNAKAGNDDIFVGGQKMNIDG 604 |
| Db | 543 | KITDGAASSTFDLTNVVQRIEGLDNAGNVTKTKETKIIAKLGBGDDNVFVSGTTEIDG 602 |
| Qy | 605 | GDGHDVRYKDGFGNITVDTGTEAGSYTVNRKVARGDIYHEVVKRQETKVKRTET 664 |
| Db | 603 | GEYDVRHYSR-GNYGALTIDATKETEQQSYTVNRFVETGKALHEVTSHTALVGRNEK 661 |
| Qy | 665 | IQYRDYELRVKGYQSTDNLSVEEVIGSQFNDFKSGKFNDFHSGEGDLDLGGAGD 724 |
| Db | 662 | IEYR-HSNNQHHAGYITKTLKAVEEILCTSHNDIFKSGKFNDFHSGEGDLDLGGAGD 720 |
| Qy | 725 | DRLFGKGNDRLSGDEGDDLLDGGSDVNLGGAGNDVYIFRKGDGNDLTYDGTGNDKLA 784 |
| Db | 721 | DRLFGKGNDRLSGDEGDDLLDGGSDVNLGGAGNDVYIFRKGDGNDLTYDGTGNDKLA 780 |
| Qy | 785 | FADANISIMIERTEKGIIVKRNHSGSINIPRWY-----ITSNLQNTQSKTDHKIKOLI 840 |
| Db | 781 | FSDSNLKDLTPEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838 |
| Qy | 841 | GKDGSIITSDOIKLQDKGTVTISQELAKLADENKSKQLSADSLASSLNKLVGSMAL 900 |
| Db | 839 | GONGRITSKQVDDII--AKNGKITQDELKSKVDVNYELLKHS-KNVNLSLKLISVSA 895 |
| Qy | 901 | FGTANSVSNALQIPTQPTQGI 922 |
| Db | 896 | FTSSNDSRNVLVAPTSMLDQSL 917 |
| RESULT 9 | | |
| AAR42378 | | |
| ID AAR42378 | | standard; Protein; 924 AA. |
| XX | | |

| | | |
|----|---|--|
| AC | AAR42378; | |
| XX | | |
| XX | 25-MAR-2003 (updated) | |
| DT | 19-APR-1994 (first entry) | |
| DT | | |
| XX | | |
| DE | Recombinant leukotoxin peptide (split) from plasmid pGCH5. | |
| XX | | |
| XX | Haemophilus somnus; immunogenic; haemolysin; LppB; LppC; | |
| KW | thromboembolic meningencephalitis; septicaemia; arthritis; | |
| KW | pneumonia; lktA gene; haemin-binding protein; fusion protein. | |
| XX | | |
| OS | Pasteurella haemolytica. | |
| XX | | |
| PN | WO9321323-A1. | |
| XX | | |
| PD | 28-OCT-1993. | |
| XX | | |
| XX | 05-APR-1993; 93WO-CA00135. | |
| XX | | |
| PR | 09-APR-1992; 92US-0865050. | |
| PR | 04-JUN-1992; 92US-0893424. | |
| PR | 04-JUN-1992; 92US-0893426. | |
| PR | 29-MAR-1993; 93US-0038287. | |
| PR | 29-MAR-1993; 93US-0038288. | |
| PR | 29-MAR-1993; 93US-0038719. | |
| XX | | |
| XX | (UUSA-) UNIV SASKATCHEWAN. | |
| XX | | |
| PI | Harland RJ, Pfeiffer CG, Pontarollo RA, Potter AA; | |
| PI | Rioux C, Theisen M; | |
| XX | | |
| DR | WPI; 1993-351733/44. | |
| DR | N-PSDB; AAQ51081. | |
| XX | | |
| PT | Haemophilus somnus immunogenic proteins used in vaccines - | |
| PT | selected from haemin-binding protein, haemolysin, LppB and LppC, | |
| PT | and corresp. DNA | |
| XX | | |
| PS | Disclosure; Fig 5; 119pp; English. | |
| XX | | |
| CC | The hmb gene encoding the haemin-binding protein was expressed in | |
| CC | E. coli as a fusion to the Pasteurella haemolytica leukotoxin gene | |
| CC | lktA coded for by plasmid pAA352. The hmb gene fragment was taken | |
| CC | from pRAP501 and starts at the codon for the third amino acid residue | |
| CC | of ORF1. The haemin binding protein can be used in vaccines for | |
| CC | preventing or treating H. somnus infections, which cause thromboembolic | |
| CC | meningo-encephalitis, septicaemia, arthritis and pneumonia in | |
| CC | vertebrates. | |
| CC | See also AAR42370-86. | |
| CC | (Updated on 25-MAR-2003 to correct PN field.) | |
| XX | | |
| XX | Sequence 924 AA; | |
| SQ | | |
| | Query Match | 50.1%; Score 2326.5; DB 14; Length 924; |
| | Best Local Similarity | 50.0%; Pred. No. 1.7e-136; |
| | Matches | 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12; |
| QY | 15 | LNSTKSGKLNLYLAIPKD--YDPQKGTGLNDFIKAADELGIARLAEPNHTETAKKSVDT 72 |
| DB | 7 | LSFPKTKAKKIILYIPQNYQYDTQGNGLQDLVKAAEELGIEVQREERNIATAQTSLGT 66 |
| QY | 73 | VNQFLSLTGTGIAISATKLEFLQKHSNTKLAKGLDSVENIDRKLKASNVLSLSSPLG 132 |
| DB | 67 | IQTALGTERGIVLSAPQIDKLQK---TKAGALGSAESIVQNVANKAKTVLSGIQSILG 123 |
| QY | 133 | TALAGIELDSLIIKKGDAAPDALAKASIDLINEIGNLSQSTQTIEAFSSQLAKLGSTISQ 192 |
| DB | 124 | SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDFEGEQTISQFGSKLQ 182 |
| QY | 193 | AKGFSNIGNKLQNL-NFSKTNLGLIEITGLSGISAGFALADKNASTGKVAAGFELSNO 251 |
| DB | 183 | IKGLGTIGDKRNIGGLDKAGLGLDVISLGSATAALVLADKNASTAKKVGAGFELANQ 242 |

XX 24-JAN-1996; 96WO-CA00049.
 XX 10-FEB-1995; 95US-0387156.
 XX (UYSA-) UNIV SASKATCHEWAN.
 XX Manns JG, Potter AA;
 XX WPI; 1996-384447/38.
 DR N-PSDB; AAT37176.
 XX Gonadotropin-releasing hormone multimer fusion proteins - with
 PT leukotoxin polypeptide for increased immunogenicity, useful in
 PT antifertility vaccine prodn.
 XX Claim 7; Fig 5A-5H; 87pp; English.
 XX A chimeric protein (AAW03942) is composed of a fusion between
 CC a truncated leukotoxin (LKT-352) from Pasteurella haemolytica (see
 CC also AAW03945) and a 4-copy gonadotrophin-releasing hormone (GnRH)
 CC repeat sequence (see also AAW03944). It is the product of a
 CC chimeric gene (AAT37176) produced by ligating a synthetic sequence
 CC for the 4-copy GnRH into vector pAA352 (ATCC 68283), which carries
 CC the LKT-352 gene. Recombinant plasmid pCB113 (LKT 352:4 copy
 CC GnRH, ATCC 69749) was obt'd. Escherichia coli transformants
 CC produced the chimeric protein, which is useful as a vaccine for
 CC fertility control, esp. immunological sterilisation of
 CC domestic or farm animals.
 XX SQ Sequence 977 AA;

Query Match 50.1%; Score 2326.5; DB 17; Length 977;
 Best Local Similarity 50.0%; Pred. No. 1.9e-136;
 Matches 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12;

QY 15 LNSTKSGLNLYLAI PKD--YDPOKGGTLDNFIKADELGTARLAEPNHTETAKKSVDT 72
 DB 7 LSPFKTAGKILYIPQNYQVDTGGGLQDLVAASELGLVEOREERNIATQTSIGT 66
 QY 73 VNQFLSTQTGIAISATKLEKFLQKSTNKLAKGLSDSVENIDRKLKASNVLSLTSFLG 132
 DB 67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNAKAKTVLSGIIQILG 123
 QY 133 TALAGIELDSLKGGDAAPDALAKASDLINIEIIGLSQSQTOTIEAPSSQLAKLGSITSQ 192
 DB 124 SVLAGMDLDEAL-QNNSNQHALAKAGLELNTSLTENTANSVKTLDDEFGEQISQKLN 182
 QY 193 AKGFSNIGKLNQNL-NFSKTNLGLIEITGLLSGISAGFALADKNAAGTKKVAAGFELSNO 251
 DB 183 IKGLGTIGDKUKNIGGLDKAGLGLDVLGSLSGATNALVLADKNAATKAKGAGFELANQ 242
 QY 252 VIGNVTKAISYVLAQRVAAGLSTTGVAALITSIMLAISPLAFMNAADKFNHANALDE 311
 DB 243 VVGNITRAVSVYLAQRVAAGLSSTGPPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302
 QY 312 FAKQFRKFGDGHLLAYQKRGVGTIEASLTITSTALCAVSAGSAAVSAVGTPIALL 371
 DB 303 YAEKFKLGLGDGNLLAEYQKRGVGTIDASVTAINTALAAIAGGVSAAGSVIASPIALL 362
 QY 372 VAGVTGLISGLEASKQAFESVANRQKLEWKEKONGONYFDKGVDYSYAYLANNL 431
 DB 363 VSGTGVISTITLQVSKQAFHVAANKHINKIVVEKKNHGNKYPENGDYARLANLQNM 422
 QY 432 KFLSELNKEBAERVAITQQRWNNIGELAGITKLGRIKSGKAYADAFEDGKKVEAGS 491
 DB 423 KFLNKLNLQERVAITQQWNNIGLGLAGISPLGEKVLSGKAYDAFEGKHAKD 482
 QY 492 NITLDKAGTIIDINSNGKKTQALHFTSPLITAGTESRERTNGKYSYINKLKFGVRQNW 551
 DB 483 LVQLDSANGIIDVNSNGKAKTQHILFRPLTTPGTEHRRVQTGKYEYITKLNINRVDSW 542
 QY 552 QVTDEAASSKLDFFSKVIQRV-----AETGTDEIGLIVNAKAGNDDIFVQGGKMNIDG 604

DB 543 KITDGAASSTFDLTNNVQRIEGLDNAGNVTKTKETKIAKLGEGDNNVFGSGTTEIDG 602
 QY 605 GDGHRVFTSKGGPGNITVDGTSATFAGSYTVNRKVARGDIYHEVVKROETKVGKRTET 664
 DB 603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVSTHTALTALVGNREK 661
 QY 665 IOYRDYELRKVGYGYSOTNLKSVFVIGSQFNDVFKSGKFNDIFHSGEGLDLDGGAGD 724
 DB 662 IEYR-HSNQOHAGYVTKTLKAVEELIIGTSHNDIFKSGKFNDAFNGGDSVDVTDGNDGN 720
 QY 725 DRLFEGKGNDRLSGDEGDDLLDGGSGDVLNGAGNDVYIFRKGNDNTLYDGTGNDKLA 784
 DB 721 DRLFEGKGGDDILDGGNGDDFDGKGNDLLHGKGGDDIFVHRKGGDNDIITDSGNDKLS 780
 QY 785 PADANISDIEMIERTKEGIIVKRNHSGSINIPWY----ITSNLQNVQSNKTHKIEQLI 840
 DB 781 FSDNLKDLTTFEKVKNLVI-TNSKKEKVTIQWFREDAFKEVPNPKATK-DEKIEEII 838
 QY 841 KGDSYITSDQIDKILQDKKDGTVITSOELKLADENKOKLSASDIASSLNKLVGSMAL 900
 DB 839 GQNGERITSKQVDDLI--AKNGKITQDELKSKVDNYVELLKHS-KVNTNSLDKLISVSA 895
 QY 901 FGTANSVSSNALQPIITQPTQGI 922
 DB 896 FTSSNDSNRNVLVAPTSMLDQSL 917

RESULT 15
 AAW79569
 ID AAW79569 standard; Protein; 977 AA.
 XX AC AAW79569;
 XX DT 24-DEC-1998 (first entry)
 XX DE LKT-GnRH chimeric protein.
 XX KW Chimera; pCB113; LKT 352; GnRH; Gonadotropin releasing hormone; multimer;
 KW cytotoxic activity; antigen presentation; immune response; vaccine;
 XX tumour.
 OS Synthetic.
 XX PW0906848-A1.
 XX PD 19-FEB-1998.
 XX PF 08-AUG-1997; 97WO-CA00559.
 XX PR 09-AUG-1996; 96US-0694865.
 XX (UYSA-) UNIV SASKATCHEWAN.
 XX Manns JG, Potter AA;
 XX WPI; 1998-159540/14.
 DR N-PSDB; AAV61531.
 XX Chimeric protein of leukotoxin and gonadotropin releasing hormone -
 PT useful for, e.g. preparation of vaccines for reduction of incidence
 PT of mammary tumours in mammals
 XX Disclosure; Figure 5.1-8; 118pp; English.
 XX The present sequence represents the LKT-GnRH chimeric protein from
 CC pCB113. This plasmid contains the LKT 352 polypeptide (AAW79568) fused
 CC to four copies of the GnRH peptide. This chimera lacks cytotoxic
 CC activity which enables there to be an increase in antigen presentation
 CC and thus an optimal immune response. The removal of this region also
 CC enables the truncated LKT to be expressed at much higher levels and
 CC allows the amount of antigen administered to be reduced. This chimeric
 CC protein comprises a leukotoxin polypeptide, several multimers, and a GnRH

QY 193 AKGFSNIGNKLNLF-NFSKTNLGLLEIITGLSGISAGFALADKNASTGKVAAGFELSNO 251
 DB 183 IKGLGTGDKLNKNGGLDGLKAGLGDVLSGLSGATAALVLADKNASTAKKVGAGFELANQ 242
 QY 252 VIGNVTKAISYVLAQRVAAGLSTTGAAALITSSIMLAISPLAFMNAADKFNHANALDE 311
 DB 243 VVGNITKAVSSYILAQRVAAGLSTGPAVALIASTVSLAISPLAFAGIADKFNHAKSLES 302
 QY 312 FAKQFRKFGYDGDHLLAEYQYRGVGTTEASLTSTTALGAVSAGSAAAVGSAVGTPIALL 371
 DB 303 YAEFRKGLGYDGNLLAEYQYRGVGTTEASLTSTTALGAVSAGSAAAVGSAVGTPIALL 362
 QY 372 VAGVTGLISGLLEASKAMPESVANRLOGKILEWKEKONGYDFKGVDSRYAAYLANNL 431
 DB 363 VSGITGVISTILQYSKOAMFEHANKHIVKWKNNHGNKYFENGDIADKLANLQDNM 422
 QY 432 KFLSELNKELEAERVIAITQQRWNNIGELAGITKGLERIKSGKAYADAFEDGKKEVAGS 491
 DB 423 KFLANLNLKELQAERVIAITQQRWNNIGELAGISRLGKVLGKAYYDAFEEGKHAKD 482
 QY 492 NITLDATGIIIDISNGKKTQALHFTSPLLTAGTESRERLUTNGKYSYINKLKGPRVQNW 551
 DB 483 LVQLDANGIIVDVSNGKAKTQHILFRLPLTPGTEHREVRQTKYBYITKLNINRVDSW 542
 QY 552 QVTDGEASSKLDKFSKVIQV-----AETGTDIEGLIVNAKAGNDDIFVGGCKMNDG 604
 DB 543 KITDGAASSTFDLTNVVQRIEILDNAGNVTKTKETKIIAKLGGEDNVFVGSSTTEIDG 602
 QY 605 GDGHRVYFSKDGFGNITVDGTSATBAGSYTVNRKVARGDYIEHVVKROETKVGKRTET 664
 DB 603 GEGYDRVHYSR-GNYGALTIDATKETEGSYTVNRVETGKALHEVTSHTALVGNREK 661
 QY 665 IQYRDYELRVKGYQSTDNLSKVEEIVGSGFNDVFKGSKFNDIFPHSGEGDLDLDGAGD 724
 DB 662 IEYR-HSNQNHAGYTTKTLKAVEEITGTSNDIFKGSKFNDAFNGGSDGVDTTIDGNDGN 720
 QY 725 DRLFGGKGNDRLSGDEGDDLLDGGSDVNLGAGAGNDVYIFRKGDDGNDLYDGTGNDKLA 784
 DB 721 DRLFGGKGDLDGGGDDPFDGGKGNLDLHGGKGGDDIFVHRKGGDNDITDSDGNKLS 780
 QY 785 FADANISDIIMERTKEGIVKRNDSHGSINIPRY-----ITSNLQNTQSNKTDHKBQLI 840
 DB 781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFRADFAKEVPNYKATK-DEKIEEII 838
 QY 841 GKGSYITSDQIDKIQKXKDGTVITSOELKKLADENKSKLSASDIASSLUNKLVGSNAL 900
 DB 839 GONGERITSKQVDDLI--AKNGKITQDELKSVVDNVELLKHS-KNVTNSLDKLISSVSA 895
 QY 901 FGTANSVSSNALQIPQTQGI 922
 DB 896 FTSSNDSRNLVAPTSMLDQSL 917
 RESULT 17
 AAW13867
 ID AAW13867 standard; Protein; 1069 AA.
 AC AAW13867;
 XX
 XX 25-MAR-2003 (updated)
 DT 12-MAY-1997 (first entry)
 XX
 XX Chimeric protein #2.
 DE
 XX RTX cytotoxin; cytokine; immunogen; chimeric protein; cytokine; vaccine;
 KW interleukin-2; IL-2; gamma interferon; gamma IFN; leukotoxin; pneumonia;
 KW Pasteurella haemolytica; LKT352; respiratory disease; shipping fever;
 KW fibrinous pneumonia; cattle; therapy.
 XX
 OS Synthetic.
 XX
 XX US5594107-A.
 PN

XX 14-JAN-1997.
 PD
 XX 20-DEC-1993; 93US-0170126.
 PF
 XX 20-DEC-1993; 93US-0170126.
 PR 22-AUG-1990; 90US-0571301.
 PR 16-OCT-1991; 91US-077715.
 XX (CIBA) CIBA GEIGY CANADA LTD.
 PA (UYSA-) UNIV SASKATCHEWAN.
 XX Campos M, Hughes HPA, Potter A;
 PI
 XX WPI; 1997-099529/09.
 DR N-PSDB; AAT60033.
 XX
 PT Immunogenic chimeric proteins comprising cytokine linked to RTX
 PT toxin - useful in vaccines, esp. against shipping fever in cattle
 XX
 PS Claim 13; Column 37-46; 56pp; English.
 XX AAW13866 and AAW13867 represent immunogenic chimeric proteins of the
 CC invention. This sequence represents a chimeric protein containing the
 CC bovine gamma interferon (gamma IFN) sequence and a leukotoxin sequence.
 CC The chimeric proteins of the invention comprise a cytokine, selected
 CC from interleukin-2 (IL-2) and gamma IFN, linked to at least one RTX
 CC toxin epitope (preferably the sequence shown in AAW13865). The RTX toxin
 CC used to provide the epitope sequence is preferably a leukotoxin,
 CC especially the full-length Pasteurella haemolytica leukotoxin.
 CC Alternatively, the leukotoxin is a truncated leukotoxin lacking
 CC leukotoxic activity, especially LKT352. The chimeric proteins can be
 CC used for the production of vaccines against respiratory diseases such as
 CC pneumonia, particularly fibrinous pneumonia caused by P.haemolytica,
 CC including shipping fever in cattle.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 1069 AA;
 Query Match 50.1%; Score 2326.5; DB 18; Length 1069;
 Best Local Similarity 50.0%; Pred. No. 2.1e-136;
 Matches 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12;
 QY 15 LNSTKGLKLNLYAIKPD--YDPQKGTGTLNDFKAADELGIARLAEPNHTETAKKSVDT 72
 DB 7 LSPFKTAKKIIYIPONYOYDFEQNGLODLVAAAEELGIEVQREERNRIATAQTSLGT 66
 QY 73 VNQFLSLTQTGIAISATKLEKFLQKISTNKLAKGLDSVENIDRKLKASNVLSLSSFLG 132
 DB 67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQILG 123
 QY 133 TALAGIELDSLKKGDAAADALAKASIDLNEIIGLSQSTQTEAFSSQLAKLGSTISQ 192
 DB 124 SVLAGMDLDEAL-QNNSNQHAKAKAGLELTNLSLTENIANSVKTLDTEGEQISQFSGKLQ 182
 QY 193 AKGFSNIGNKLNLF-NFSKTNLGLLEIITGLSGISAGFALADKNASTGKVAAGFELSNO 251
 DB 183 IKGLGTGDKLNKNGGLDGLKAGLGDVLSGLSGATAALVLADKNASTAKKVGAGFELANQ 242
 QY 252 VIGNVTKAISYVLAQRVAAGLSTTGAAALITSSIMLAISPLAFMNAADKFNHANALDE 311
 DB 243 VVGNITKAVSSYILAQRVAAGLSTGPAVALIASTVSLAISPLAFAGIADKFNHAKSLES 302
 QY 312 FAKQFRKFGYDGDHLLAEYQYRGVGTTEASLTSTTALGAVSAGSAAAVGSAVGTPIALL 371
 DB 303 YAEFRKGLGYDGNLLAEYQYRGVGTTEASLTSTTALGAVSAGSAAAVGSAVGTPIALL 362
 QY 372 VAGVTGLISGLLEASKAMPESVANRLOGKILEWKEKONGYDFKGVDSRYAAYLANNL 431
 DB 363 VSGITGVISTILQYSKOAMFEHANKHIVKWKNNHGNKYFENGDIADKLANLQDNM 422
 QY 432 KFLSELNKELEAERVIAITQQRWNNIGELAGITKGLERIKSGKAYADAFEDGKKEVAGS 491
 DB 423 KFLANLNLKELQAERVIAITQQRWNNIGELAGISRLGKVLGKAYYDAFEEGKHAKD 482

Db 423 KFLNLNKLQAEVIAITQQQWNNNIGDLAGISRLGKVLGSKAYVADPEEGKHAKD 482
 Qy 492 NITLDKTIIDISNSNGKTKQALHFTSPPLTAGTESRERLTNGKYSYINKLKFGVRVNW 551
 Db 483 LVQLDSANGIIDVNSGKAKTQHILFRTPLTPGTEHRRVQTGKYEYITKLNINRVDSW 542
 Qy 552 QVTGEASSKLDKFSKIVRV-----AETGTEIGLIVNAKAGNDDIFVGGCKMIDG 604
 Db 543 KITDGAASSTFDLTNNVQRIEILDNAGNVTKTKETIIAKLGEEDDNVFGSGTTEIDG 602
 Qy 605 GDGHRVFPYSKDGFGNITVDTGTSATEAGSYTVNRKVARGDYIHEVVKRQETKVKRTET 664
 Db 603 GEGDVRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSHTALVGNREEK 661
 Qy 665 IQYRDYELRKVGYSQSDNLKSVVEVIGSQFNFVFGSKFNDIFHSGEGDLDLGGAGD 724
 Db 662 IEYR-HSNNOHQHAGYTKDTLKAVEEIICTSHNDIFKSGKFNDAFNGGDDVTIDGNDG 720
 Qy 725 DRLFGKGNDRLSGDEGDDLDGGSDVNLGGAGNDVYIERKGGNDLTDGTGNDKLA 784
 Db 721 DRLFGKGGDDLDGGNGDDFDGKGKNDLHGGKGGDDIFVHRKGGDNDIITDSGNDKLS 780
 Qy 785 FADANISIMIERTEKEGIIVRNDSHSGSINIPRWY----ITSNLQNYQSNKTDHKIEOLI 840
 Db 781 FSDSNLKDLTETKVKHNLVI-TNSCKEVTIQNWFEADFAKEVPNYKATK-DEKIEEII 838
 Qy 841 KGQSYITSDQIDKTLQDKGTVITSOELKKLADENKSKLSASDTSASSLNKLVGSMAL 900
 Db 839 GQNGERITSKQVDDLI--AKNGKITQDELKSVVDNYELLKHS-KNVNTSLDKLISVSA 895
 Qy 901 FGTANSVSSNALQPIQTQGI 922
 Db 896 FTSSNDSRNLVAPTSMLDQSL 917

RESULT 18
 AAB21074
 ID AAB21074 standard; Protein; 1069 AA.
 AC AAB21074;
 XX AAB21074;
 DT 19-DEC-2000 (first entry)
 XX Bovine gamma-IFN/Pasteurella haemolytica leukotoxin fusion gene.
 DE Bovine gamma-IFN; gamma-interferon; leukotoxin; LKT; respiratory disease;
 KW pneumonia; shipping fever; cattle; livestock; anti-Pasteurella vaccine;
 XX immunogen.
 OS Chimeric - Bos taurus.
 OS Chimeric - Pasteurella haemolytica.
 XX US6096320-A.
 PN 01-AUG-2000.
 PD 20-OCT-1997; 97US-0954418.
 PF 20-DEC-1993; 93US-0170126.
 PR 22-JUL-1996; 96US-0681479.
 PR 22-AUG-1990; 90US-0571301.
 PR 16-OCT-1991; 91US-0777715.
 XX (UYSA-) UNIV SASKATCHEWAN.
 PA (CIBA) CIBA GEIGY CANADA LTD.
 XX Campos M, Hughes HPA, Potter A;
 XX WPI; 2000-531543/48.
 DR N-PSDB; AA72484.
 XX Vaccine for stimulating immunity against pneumonia comprises chimeric
 PT protein comprising gamma-interferon and leukotoxin derived from

PT Pasteurella haemolytica -
 XX Claim 8; Column 45-50; 56pp; English.
 XX This sequence a fusion protein comprising bovine gamma-interferon
 CC (gamma-IFN) and Pasteurella haemolytica leukotoxin (LKT). The
 CC fusion protein is immunogenic, and may be used in an anti-
 CC Pasteurella vaccine composition. Pasteurella species, especially
 CC Pasteurella haemolytica, are responsible for respiratory diseases
 CC in a range of agricultural animals, most particularly cattle, but also
 CC sheep, pigs, horses and fowl. Shipping fever is the most economically
 CC important respiratory disease associated with Pasteurella species,
 CC affecting 15-30% of exposed cattle and resulting in a 2-5% mortality
 CC rate in the exposed population. The vaccine composition of the
 CC invention is useful for preventing or ameliorating respiratory
 CC diseases such as pneumonia, particularly shipping fever pneumonia, in
 CC livestock.
 XX SQ Sequence 1069 AA;
 Query Match 50.1%; Score 2326.5; DB 21; Length 1069;
 Best Local Similarity 50.0%; Pred. No. 2.1e-136;
 Matches 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12;
 Qy 15 LNSTKSLGNLXLAIPKD--YDPQGGTLDNDFKADELGIARLAEEPNTHTAKSVDT 72
 Db 7 LSPFKTAKLILYIPQYQYDTEQNGLODLVKAEEELGIEVQREERNIATQTSIGT 66
 Qy 73 VNQPLSTQTGIAISATKLEFKLQKHSNTNKLAKGLDSVENDIRKLGKASNVLTSLSSFLG 132
 Db 67 IQTAIGLTERGIVLSAPQIDKLLQK--TRQAGALGSAESIVQNAKAKTVLSGIQSTLG 123
 Qy 133 TALAGIELSLIKGDAAPALAKASIDLINEIGNLSOSTQTIEAFSSQLAKGSTISQ 192
 Db 124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENANSVKTLDETEGEQSGSKLQN 182
 Qy 193 AKGFSNIGKNLQNL-NFSKTNLGLIEITGLISGISAGFALADKNASTKCKVAAGFELSNO 251
 Db 183 IKGLTGDKLKNIGGLDKAGLDLVISGLSGATAALVLADKNASTAKKVGAGFELANQ 242
 Qy 252 VIGNVTKAISYVLAQRAAGLSTTGAAVAALITSSIMLAISPLAFMAADKFNHANALDE 311
 Db 243 VVGNITKAVSYILAQRVAAGLSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLES 302
 Qy 312 FAKQFRKFGVDGHLAEYQGVGTIEASLTITSTALGAVSAGVSAAGVSAVGTPIALL 371
 Db 303 YAEFRKGLGDGNDLLAEYQRTGTIDASVTAINALAAIAGGVSAAGVSAVGTPIALL 362
 Qy 372 VAGVTGLISGLEASKQAMFESVANRLQKLEWEKONGQNGYFDKGYDSRYAAYLANNL 431
 Db 363 VSGITGVISTILOYSKQAMFEHVANKIHNKIVEKKNHKGKYPENGVDARYLANLQDNM 422
 Qy 432 KFLSELNKELEAEARVIAITQORWNNITGELAGITKLERIKSGKAYADAFEDGKKVEAGS 491
 Db 423 KFLNLNKLQAEVIAITQQQWNNNIGDLAGISRLGKVLGSKAYVADPEEGKHAKD 482
 Qy 492 NITLDKTIIDISNSNGKTKQALHFTSPPLTAGTESRERLTNGKYSYINKLKFGVRVNW 551
 Db 483 LVQLDSANGIIDVNSGKAKTQHILFRTPLTPGTEHRRVQTGKYEYITKLNINRVDSW 542
 Qy 552 QVTGEASSKLDKFSKIVRV-----AETGTEIGLIVNAKAGNDDIFVGGCKMIDG 604
 Db 543 KITDGAASSTFDLTNNVQRIEILDNAGNVTKTKETIIAKLGEEDDNVFGSGTTEIDG 602
 Qy 605 GDGHRVFPYSKDGFGNITVDTGTSATEAGSYTVNRKVARGDYIHEVVKRQETKVKRTET 664
 Db 603 GEGDVRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSHTALVGNREEK 661
 Qy 665 IQYRDYELRKVGYSQSDNLKSVVEVIGSQFNFVFGSKFNDIFHSGEGDLDLGGAGD 724
 Db 662 IEYR-HSNNOHQHAGYTKDTLKAVEEIICTSHNDIFKSGKFNDAFNGGDDVTIDGNDG 720
 Qy 725 DRLFGKGNDRLSGDEGDDLDGGSDVNLGGAGNDVYIERKGGNDLTDGTGNDKLA 784

```

Db 721 DRLEGGGDDILDGNGDDPFDGKGNLHGGKGGDI FVHRKGGDNDIITDSGNDKLS 780
Qy 785 PADANISDIEMIETKEGIIKVRNDHSGSINI PRWY-----ITSNLQNTQSKNTHKIEQLI 840
Db 781 FSDSNLKDLPFKVKHNLVI--TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838
Qy 841 GKDSYITSDOIKILQDKDGTVTISQELKKLADENKSKLSASDIASSINKLVGSMAL 900
Db 839 GQNGERITSQVDDLI--AKNGKIIQDELKSKVDNYELLKHS-KNVTNSLDDKLISSVA 895
Qy 901 FGTSVSSNALQPIPTQGI 922
Db 896 FTSSNDSRNVLVAPTSMLDQSL 917

RESULT 19
AAR07167
ID AAR07167 standard; protein; 953 AA.
XX AC AAR07167;
XX DT 24-JAN-1991 (first entry)
XX DE 105kD PTX protein of P.haemolytica.
XX KW Vaccine; shipping fever.
XX OS Pasteurella haemolytica.
XX PN US4957739-A.
XX PD 18-SEP-1990.
XX PF 13-AUG-1987; 87US-0085430.
XX PR 13-AUG-1987; 87US-0085430.
XX PA (TEXA ) UNIV OF TEXAS SYSTE.
XX PI Berget P, Engler M, Highlander S, Weinstock G;
XX DR WPI; 1990-304558/40.
XX DR N-PSDB; RAQ06074.
XX PT Vaccine against shipping fever in cattle - contains purified
XX PT Pasteurella haemolytica antigen of molecular wt. 105 kD
XX PS Disclosure; Fig 9a-h; 35pp; English.
XX CC Gene product is antigenic to P.haemolytica, and may be used as a
XX CC vaccine for immunisation against shipping fever. Abs raised to the
XX CC antigen may be used in passive immunisation and diagnosis.
SQ Sequence 953 AA;

Query Match 50.0%; Score 2325; DB 11; Length 953;
Best Local Similarity 49.9%; Pred. No. 2.2e-136;
Matches 464; Conservative 175; Mismatches 264; Indels 26; Gaps 13;

Qy 8 KSNIOAGLNSTKGLKNLYLAI PKD--YDPQKGGTLDNDFIKAADELGIARLAEPNHTET 65
Db 28 QSLTQAG-SSUKTAGKIIIPQNYDTQGGQLQDLVKAASELGLIEVQREERNIAT 86
Qy 66 AKKSYDVTNQFLSTQTGIAISATKLEKFLQKHSTKNLAKGLSDSVENIDRKLGSNVLS 125
Db 87 AQTSLGTITQALGTERGIVLSAPQIDKLQK---TKAGQALGSAESIVQVANKAKTVLS 143
Qy 126 TLLSFLGTALAGIELDSLIKGGDAPDALAKASIDLINIEIGNLSQSQTQTIEAFSSQLAK 185
Db 144 GIQSTLGSVLGMDLDEAL-QNNSNQHALAKAGLELTNLSLIENIANSVKTLDFEFGQISQ 202
Qy 186 LGSTISQAKGFSNIGNKQLNL-NFSKTNLGLLEIITGLLSGISAGFALADKNASTGKKVAA 244

```

```

Db 203 FGSKLQNIKGLTGLDKLNIGGLDRAGLGLDVISGLLSGATAALVLADKNASTAKKVG 262
Qy 245 GPELSNQVIGNVTKALISSVLAQRAAGLSTTCGAAALITSSIMLAISPLAFNADAKFN 304
Db 263 GPELANQVGNITKAVSSYILAQRVAAGLSSTGPFVAALITASTVSLAISPLAFAGIADKFN 322
Qy 305 HANALDEFAPKOPKFGVGDHLLAEYQORGVTIEASLTTISTALGAVSAGVSAAGVSAV 364
Db 323 HAKLSYAEERFKLGYDGDNLAEYQRGTTIDASVTAINTALAAIAGVSAAGVSAV 382
Qy 365 GTPIALLVAGVTGLISGLEASQAMPFESVANLQOKILEWEKONGQONTYDFKGYDSRYA 424
Db 383 ASPIALLVSGITGVI STILQYSQAMPFHVANKIHKNKIVEWENKNGHKNYFENGIDARYL 442
Qy 425 AYLANNLKFLSELNKELEAERVIAITQORWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db 443 ANLQDNMKFLNLNKLQAEVIAITQQQDNNIGLAGISRLGEKVLGSKAYADAFEEG 502
Qy 485 KQVEAGSNITLDAKTGIIDISNSNGKTOALHFTSPLLTAGTESRRLTNGKYSYNKLK 544
Db 503 KHIKADKLVOLDSANGIIDVNSGKAKTQHILFPTPLTPTGTEHRRVOTGKYEITKLN 562
Qy 545 FGRVKWQVTDGEASSKLDPSKVIQVY-----AETEGTDEIGLIVNAKAGNDDIPVQG 597
Db 563 INRVDSWKITDGAASSTFDLTNVVQRIEGLDNGVNTKTKETIIAKLGEGLDNNVFGS 622
Qy 598 GKWNIDGGDHRVYFSKGGFGNITVDGTSATEAGSYTVNRKVARGDIIYHEVVKQETK 657
Db 623 GTTEIDGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVYTHAL 681
Qy 658 VGRKTETIOYRDELKRVGYGYOSTDNLKSVBEVIGSQFNDVPKSKFNDFHSGEGDDL 717
Db 682 VGNREKIEYR-HSNQNHAGYTKYLKAVEEIIIGTSHNDIPKGSKFNDAFNGGVDVT 740
Qy 718 LDGAGDRLFGGKGNDRLSGDEGDDLDDGGSGDDVLNGGAGNDVYIFRKGDNNTLYDG 777
Db 741 IYNGDGNDRLFGGKGGDDILDGGNGDDFIDGGKGNLHGGKGGDIFVHRKGGDNDIITDS 800
Qy 778 TGNDKLAFADANISDIEMIETKEGIIKVRNDHSGSINI PRWY-----ITSNLQNTQSKNTH 833
Db 801 DGNDKLSFSDSNLKDLPFKVKHNLVI--TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 858
Qy 834 HKIEQLIGKDSYITSDOIKILQDKDGTVTISQELKKLADENKSKLSASDIASSINK 893
Db 859 EKIEEIIQNGERITSQVDDLI--AKNGKIIQDELKSKVDNYELLKHS-KNVTNSLDDK 915
Qy 894 LVGSMALFGTSVSSNALQPIPTQGI 922
Db 916 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 944

RESULT 20
AAR60072
ID AAR60072 standard; protein; 953 AA.
XX AC AAR60072;
XX DT 25-MAR-2003 (updated)
XX DT 13-MAR-1995 (first entry)
XX KW PtxA protein of Pasteurella haemolytica.
XX OS Pasteurella haemolytica.
XX PN US5336491-A.
XX PD 09-AUG-1994.

```



```
PF 15-JUN-1992; 92US-0899100.
XX
PR 28-NOV-1986; 86US-0935806.
PR 13-AUG-1987; 87US-0085430.
PR 18-JUN-1990; 90US-0540261.
PR 15-JUN-1992; 92US-0899100.
XX
XX (TEXA ) UNIV TEXAS SYSTEM.
XX
PI Berget P, Engler M, Highlander S, Weinstock G;
XX
XX WPI; 1994-255245/31.
XX
XX N-PSDB; AAQ70050.
XX
XX New purified antigen from Pasteurella haemolytica - useful in
XX vaccines against shipping fever of cattle, also for raising
XX antibodies useful in diagnosis and passive immunisation
XX
XX Disclosure; Figure 9; 35pp; English.
XX
XX This protein is a 105 kD antigen which may be used in vaccines to
XX protect cattle against shipping fever. The antigen is also useful
XX for raising antisera which can be used for passive immunisation (for
XX treatment or temporary prophylaxis) and also to raise antibodies
XX which can be used in immunoassay methods for the detection of
XX Pasteurella haemolytica antigens in biological fluids.
XX (Updated on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 953 AA;

Query Match 50.0%; Score 2325; DB 15; Length 953;
Best Local Similarity 49.9%; Pred. No. 2.2e-136;
Matches 464; Conservative 175; Mismatches 264; Indels 26; Gaps 13;

QY 8 KSNIQAGLNSTKGLKLNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEPNHPYET 65
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
28 QSLTQAG-SLTKGAKKILYIPQNYQYDTQEGNGLQDLVRAAEELGIEVQREERNIAT 86
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 66 AKKSVDTVNOQLSTQTGTATISATKLEKPLQKHSTNKLAKGLDSVENIDRKLKGNVLS 125
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 87 AQTSLGTITQATGLTERGIVLSAPOIDKLLQK---TKAQALGSAESIVQNAKAKTVLS 143
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 126 TLSSFLGTALAGIELDILIKKGAAPDALAKASIDLINEIIGNLSOSTQTTEAFSSQLAK 185
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 144 GQSLGSLVAGWDLDEAL-QNNSHALAKAGLELTNSLIENIANSVKTIDFEGEQISQ 202
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 186 LGSTTISOAKGFSGNIGKLNQL-NFSKTNLGLBIITGLLSGISAGPALADKNAKSTKKVAA 244
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 203 FGSKLNKIKGLTGLDKLNIGGLDKAGLGLDVISGLLSGATAALVLADKNAKSTAKVGA 262
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 245 GFELSNQVIGNTKAISSVLAQORVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 263 GFELANQVVGNTTKAVSSYLAQORVAAGLSSSTGPVAALIASTVSLAISPLAFAGIADKFN 322
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 305 HANALDEFAPKQFRKFGYDGDHLLAEVQGVGTIEASLTITSTALCAVSAGVSAAAAGVSAV 364
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 323 HAKSLAESVAERPKLGYDGDNLAEYQRTGTIDASTAINALAAIAGGVSAAGVSAV 382
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 365 GTFPIALLVAGVTGLSIGILEASKQAMFESVANRLQKILEWEKQNGQGVYFDKGYDSRYA 424
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 383 ASPIALLVSGITGVLTILQYSKQAMFEHVANKIHKIVEKKNHKNYPENGYDARYL 442
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 425 AYLANNLKPLSELNKELEAERVIATQQRWDNNIGELAGITKLGRIKSGKAYADAFEDG 484
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 443 ANLQDNMKFLLNLNKELOAERVIATQQQWDDNIGLAGISRLGSKVLSGKAYVDAFEEG 502
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 485 KKVEAGSNITLDAKGTIIDSNSCKTQALHFTSPLLTAGTESRERLTNGKYSVINKLK 544
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 503 KHIKADKVLQDLSANGIDVDSNGKAKTQHILFRPLTPGTEHREVRQTKYEITKLN 562
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 545 FQVRKNQVWTDGEASSKLDFSKVIOQRV-----AETGTDIGLIVNAKAGNDDIFVQG 597
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 563 INRVDSWKITDGAASSTFDLTNVQRIIGIELDNAGNVTKTKETKIIAKLGEADDNVFVGS 622
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```


KW Leukotoxin; probe; pasteurellosis; cytotoxicity; leukocytes.

OS Pasteurella haemolytica A1 (ATCC 43270).

XX US055400-A.

XX 08-OCT-1991.

XX 26-NOV-1986; 86US-0935493.

XX 26-NOV-1986; 86US-0935493.

XX (UWVU-) UNIV OF GUELPH.

XX Lo RYC, Shewen PE, Strathdee CA;

XX WPI; 1991-317648/43.

XX N-PSDB; AAQ14000.

PT DNA encoding Pasteurella haemolytica-A1 leukotoxin - used to
produce recombinant protein for prep. of vaccine and antibodies
for protection of cattle from pneumonic pasteurellosis

XX Claim 6; Page 15; 20pp; English.

CC ORP2 (AAR14223) codes for a 19,820 D protein of 166 amino acids;
CC ORP1 codes for leukotoxin, a 101,883 D protein of 952 amino acids,
CC which is produced on expression of pLKT5 in a host organism.
CC The leukotoxin has cytotoxic activity specific against leukocytes.
CC The protein is used to prepare vaccines or to raise antibodies
CC against P. haemolytica to protect cattle from pneumonic pasteurellosis.

XX Sequence 953 AA;

Query Match 49.8%; Score 2313; DB 12; Length 953;

Best Local Similarity 49.6%; Pred. No. 1.3e-135;

Matches 461; Conservative 176; Mismatches 266; Indels 26; Gaps 13;

QY 8 KSNTOAGLNTSKGLNLYLAIPKD--YDPQKGTINDFIKADELGIARLAEPNHTET 65
DB 28 QSLTQAG-SSLTKGAKIILYIPQNYQYDTQEGNGLQDLVKAABEELGIEVQREERNIAT 86
QY 66 AKKSVDTVQPLSTOTGIATISATKLEKFLQKHSTNKLAKGLDSVENIDRKLKGSANVLS 125
DB 87 AOTSIGTQTAIGTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQVANKAKTVLS 143
QY 126 TLSSFLGTALAGIBLSDLIKKGDAAPDALAKASIDLINEITGNLSQSTQTIEAFSSQLAK 185
DB 144 GIQSILGSLVAGMDLDEAL-QNNSQHALAKAGLELTNSLIENIANSVKTLDPEFGQISQ 202
QY 186 LGSTISQAKFSNTGNKQNL-NFSKTNLGLIEITGLLSGISAGFALADKNASTGKKVAA 244
DB 203 FGSKLQNIKGLTGLDKLKNIGGLDKAGLGLDVIISGLISGATAALVLADKNASTAKVGA 262
QY 245 GFELSNQVIGNVTKAISSYVLAQRAAGLSTGTAVALITSSIMLAISPLAFMNAADKEN 304
DB 263 GFELANQVIGNVTKAISSYVLAQRAAGLSTGTAVALITSSIMLAISPLAFMNAADKEN 322
QY 305 HANALDFAKQFRKFGYDGDHLLAEYQGVGTIEASLTITISTALGAVSAGVSAAGVSAV 364
DB 323 HAKSLESYAERFKKLYDGDNDLLAEYQGVGTIDASVTAINALAAIAGGVSAAGVSAV 382
QY 365 GTPALLVAGVTGLISGILEASQKMPESVANRLOGKILEWEKONGGONYPDKGYDSRYA 424
DB 383 ASPIALLVSGITGVISTTLOYSQAMFEHVANKIHNKIVWEKKNHGKNYPENGVDARYL 442
QY 425 AVLANNKLSELNKELEAERVIATQQRWNNNIGELAGITKLERIKSGRAYADAFEDG 484
DB 443 ANLQNMKFLNKLKELEAERVIATQQRWNNNIGELAGITKLERIKSGRAYADAFEDG 502
QY 485 KKVAGSNITLDKATGIIDISNSGKKTQALHFTSPLTAGTESRRLNTGKSYINKLK 544
DB 503 KHIRADKLVLQDSANGIIDVNSGKAKTQHILFRTPLLTPGTEHRERVQTCYKYEITKLN 562

QY 545 FGRVKNQVTDGEASSKLDPSKVIQRV-----AETEGTDEGLIVNAKAGNDIDFVGQ 597
DB 563 INRVDSWKITDGAASSTFDLTVVQRIEGLDAGNVTTKETIIAKLGEGLDNDVFGS 622
QY 598 GKXNIDGGDGHDRVYFYSKDGFGNITVDGTSATSEAGSYTVNRKVRGDIYHEVVYKQETK 657
DB 623 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEGQSGSYTVNRVETGKALHEVTSTHTAL 681
QY 658 VGRRTETIQRDYELRVKGYGYQSTNLSKSVIEVQSNDFVSKGFNDIFHSGEGDDL 717
DB 682 VGNREEKIEYR-HSNNQHAGYTYTKDLKAVEBIIIGTSHNDIFKSGKFNDAFNGGQVDT 740
QY 718 LDGAGGDDRLFGGKGNDRLLSGDDLLDGGSGDDVLLNGAGNDVYIFRKGDCGNDLTLYDG 777
DB 741 IDGNDGNDRLLFGGKGGDDILDGGNGDDFIDGGKGNDDLLHGGKGGDIFVHRKGGDNDITDS 800
QY 778 TGNDKLAFADANISDIMEERTKEGIIIVKRNDHSGSINIPRW----ITSNLQYQSNKTD 833
DB 801 DGNLKLFSFSDSNLKDILTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPVNYKATK-D 858
QY 834 HKIEQLIKDGSYITSDQIDKILQDKDGTIVTSQELKLDENKSKLSASDIASLNK 893
DB 859 EKIEEIIQNGERITSQVDDLI--AKGNGKITQDELSKYVDNYELLKHS-KNVTNSLDK 915
QY 894 LVGSMALFGTANSVSSNALQIPITQPTQGI 922
DB 916 LLSVSAFTSSNDSRNLVLAFTSMLDQSL 944
RESULT 27
AAR34548
ID AAR34548 standard; Protein; 951 AA.
XX AAR34548;
AC AAR34548;
XX 25-MAR-2003 (updated)
DT 23-AUG-1993 (first entry)
XX Rotavirus VP4-leukotoxin gene fusion prod.
DE Vector; LKT 352; flanking; recombinant; antigen; somatostatin;
KW gonadotropin releasing hormone; rotavirus viral protein 4;
KW carrier protein; lactation; reproduction; SRIF.
XX Synthetic.
OS
XX Key Location/Qualifiers
FH Protein 1..926
FT Peptide /note= "recombinant leukotoxin protein"
FT Peptide 927..951
XX /note= "Rotavirus VP-4"
PN WO9308290-A1.
XX 29-APR-1993.
XX 15-OCT-1992; 92WO-CA00449.
XX 16-OCT-1991; 91US-0779171.
PR 14-OCT-1992; 92US-0960932.
XX (UWVU-) UNIV SASKATCHEWAN.
XX Hughes HPA, Potter AA, Redmond MJ;
XX WPI; 1993-152482/18.
XX N-PSDB; AAQ1322.
XX Immunological carrier system with enhanced immunogenicity -
PT comprises chimeric protein comprising leuco:toxin peptide or
PT homologous protein fused to antigen esp. somatostatin or
PT gonadotropin releasing hormone

Db 299 TGPVAALIASTVALAVSPLSNVADKFKQADLIKSYSERFQKLGDCDRLAUFHRETG 358
 Qy 336 TIEASLTITSTALGAVSAGVAAAGVSGVAGTPPIALLVAGVTGLISGLEASKQAMFESVA 395
 Db 359 TIDASVTITNTALAAISGCVGAASAGSLVGPVALLVAGVTGLITITILEYSKQAMFEHVA 418
 Qy 396 NRGQKILEWEKONGQNYFKGVDSDRYAAVLANNLKELSELNKELEAERVAITQORWD 455
 Db 419 NKVHDRVIVEKHH-KNKYFEQGYDSRHLADQDNKFLINLNKELQAEVVAITQORWD 477
 Qy 456 NNIGELAGITKLGRIKSGKAYADAFEDGKVEAGSNITLDKGTIIDISNSNGKTKOAL 515
 Db 478 NOIGDLAAISRRTDKISSGKAYVDAFEGHQSYDSSVQLDNKNGIINISNTN-RKTSQV 536
 Qy 516 HFTSPLLTAGTESRRLTNGKYSINKLKFRVKNQWVTDEASKLDPSKVIQORVA--- 572
 Db 537 LFRTPLLTPGSENRERIQEGKNSYITKLHIQORVDSWTVDGASSVDFTNVVQRIAVKF 596
 Qy 573 -----ETEGTDELGLIVNAKAGNDIDFVCGKMNIDGGDGHDRVFYSKDGFGNITVD 625
 Db 597 DDAGNIESKQTK---IANLGAGNDNVFVGSSTVIDGGDGHDRVHYSR-GETGALVID 652
 Qy 626 GTSATEAGSYTVNRKVARGDYIHYEVKQETKVGKRTETIQYRDYELRKVGYGYQSTDNL 685
 Db 653 ATAETKSGSVKRVGDSKALHETIAHTQTNVGNREEKIEYR-REDDRFHTGTVTDSL 711
 Qy 686 KSVREVIGSQNDVFKSGKFNIDFHSBGDLDLGGAGDRLFGKGNDRLSGDEGDDL 745
 Db 712 KSVBEIIGSQNDIFKGSQFDVPHGGNGVDTIDGNDGDHFLFGAGGDVTDGNGN 771
 Qy 746 DGGSGDDVLCGAGNDVYIFKPGDGNLTLYGTGNDKLAFADANISIMIERTEKEGIIVK 805
 Db 772 VGGTGNIDISGKNDIYVHTGNDGNSITDSGGODKLAFSDVNLKDLTFKKVDSSLEI- 830
 Qy 806 RNDHSGSINIPRWY-----ITSNLQYQSNKTDHKEIQIGKDGSYITSDDIKILOKDD 861
 Db 831 INQKGEKVRIGNWFLEDDLASTVANYKAT-NDRKIEEIIKGGERITSEQVDKLI--KEG 887
 Qy 862 GTVITSQBLKLADENSKQKLSASDIASSLNKLVGSMALFGTANSVSN 910
 Db 888 NNQISAEALSKVANDYNTSK-DRONVNSLAKLISVSGSFTSSSDFRNN 935

RESULT 31
 ID AAY51410 standard; protein; 956 AA.
 AC AAY51410;
 XX
 XX
 DT 05-MAY-2000 (first entry)
 XX
 DE A. pleuropneumoniae clyIIA protein.
 XX
 KW RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
 KW antiinflammatory; antiarthritic; antiabortive; treatment; pneumonia;
 KW pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
 KW shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
 KW urinary infection; peritonitis; meningitis; gastroenteritis;
 KW passive immunization; clyIIA.
 XX
 OS Actinobacillus pleuropneumoniae.
 XX
 FN US6019984-A.
 XX
 XX
 PD 01-FEB-2000.
 XX
 PF 23-DEC-1996; 96US-0772270.
 XX
 PR 01-MAR-1995; 95US-0396244.
 XX
 PA (UYGU-) UNIV GUELPH.
 XX

PI Mallard B, Rosendal S, MacInnes J, Ricciatti P;
 XX WPI; 2000-146864/13.
 DR N-PSDB; AAZ88586.
 XX
 PT Bacterial preparation comprising microorganisms which produce a member
 of the Repeats in Toxins (RTX) family, useful for treating swine
 PT pleuropneumonia, arthritis in swine, shipping fever and abortion in
 PT cattle, and sleepy foal disease -
 XX
 PS Disclosure; Column 67-72; 96pp; English.
 XX
 CC This invention describes a novel bacterial preparation (I) which
 CC comprises one or more isolated and purified strain(s) of a microorganism,
 CC cultured in tryptone yeast extract (TYE) broth, which produces one or
 CC more RTX toxins (belonging to the family of toxins referred to as Repeats
 CC in Toxins), where the strain(s) have at least one RTX toxin which is
 CC cell-associated. The products of the invention have immunostimulatory,
 CC antimicrobial, antiinflammatory, antiarthritic and antiabortive activity.
 CC The bacterial preparation may be used as vaccines for the prophylaxis and
 CC treatment of infectious diseases caused by strains of microorganisms
 CC which produce one or more RTX toxins. The infectious diseases are swine
 CC pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
 CC septicemia, nephritis, endocarditis and arthritis in piglets; shipping
 CC fever and abortion in cattle; whooping cough, sleepy foal disease or
 CC joint ill (purulent nephritis, arthritis) in foals; septicemia,
 CC polyarthritis and abortion in horses; and urinary infections,
 CC peritonitis, meningitis, and gastroenteritis. The bacterial preparations
 CC may also be used to prepare antibodies which may be used as a means of
 CC passive immunization. This sequence represents the Actinobacillus
 CC pleuropneumoniae clyIIA protein described in the method of the invention.
 XX
 SQ Sequence 956 AA;
 Query Match 48.1%; Score 2236.5; DB 21; Length 956;
 Best Local Similarity 47.7%; Pred. No. 7.4e-131;
 Matches 453; Conservative 188; Mismatches 255; Indels 53; Gaps 15;
 Qy 1 MSNINV--IKSNTQAGL-----NSTKSLGNLYLAIPKVDPPQ 36
 Db 1 MSKTLSLSKSLQQGLKNGKLNQAGTTLKNGLTQTGSHSLQNGAKKLLIYIPQGYDSG 60
 Qy 37 KGGLTNDPFIKADELGIARLAEPNHTETAKKSVDTVNFQSLTQTGTATISATKLEKFLQ 96
 Db 61 QGNGVQLVAANDLGIENVREERSNLDIAKTSPTQKILGFTDRIIVLFPAPOLDNLK 120
 Qy 97 KHSTNKLAKGLDSVENIDRLKASNVLSLTSSGFLGTALAGIELDSLIKGDAAADALAK 156
 Db 121 KNP--KIGNTLGSASSISQNIKANTVLCGIGTILGSVLSGVNLNELLQNKDPNQLLEAK 178
 Qy 157 ASDLINEIIGNLSQSTOTTEAFSSQLAKGLSTISOAKGFSNIGNKLNQL-NFSKTNLGL 215
 Db 179 AGLELTNVLGNVTASSVQTVDAFAEQISKLGSLHQLNVKGLGSLNKLQNLPLDGKASLGL 238
 Qy 216 EITLGLSGISAGFALADKNAATGCKVAAGFELSNOVIGNVTKAISSVYLAQRAAGLST 275
 Db 239 DIISGLSGASAGILLADKEASTEKKAAGVAFANOIIGNVTKAIVSYLAQRAVAGSLSS 298
 Qy 276 TGAAVALITSIMLAISPLAFMNAADKFNHANALDEPAKQFRKFGYDGDHLLASYQRCVG 335
 Db 299 TGPVAALIASTVALAVSPLSNVADKFKQADLIKSYSERFQKLGDCDRLAUFHRETG 358
 Qy 336 TIEASLTITSTALGAVSAGVAAAGVSGVAGTPPIALLVAGVTGLISGLEASKQAMFESVA 395
 Db 359 TIDASVTITNTALAAISGCVGAASAGSLVGPVALLVAGVTGLITITILEYSKQAMFEHVA 418
 Qy 396 NRGQKILEWEKONGQNYFKGVDSDRYAAVLANNLKFLSELNKELEAERVAITQORWD 455
 Db 419 NKVHDRVIVEKHH-KNKYFEQGYDSRHLADQDNKFLINLNKELQAEVVAITQORWD 477
 Qy 456 NNIGELAGITKLGRIKSGKAYADAFEDGKVEAGSNITLDKGTIIDISNSNGKTKOAL 515
 Db 478 NOIGDLAAISRRTDKISSGKAYVDAFEGHQSYDSSVQLDNKNGIINISNTN-RKTSQV 536

QY 516 HFTSPLLTAGTESRRLTNGKYSYINKLKFRGKVMQVTDGEASSKLDKFSKVIORVA--- 572
 DB 537 LFRTPLLTPGEBNRERIOEGKNSYITKLHIQKVDSWTVDGDSVDFTNVQRIAVKF 596
 QY 573 -----ETEGTDEIGLVNAKAGNDDIFVGQCKNIDGSDGHDHDFVFSKDGFGNITVD 625
 DB 597 DDAGNIIESKOTK---IIANLGNNDNVFVSSSTVVDGGDGHDRVHYSR-GEVGAVID 652
 QY 626 QTSATEAGSYTVNRKVARGDYIHEVVRQETKVGKRTETIQVRDYELAKVGYGQSTDNL 685
 DB 653 ATAETEGKSYSKRVYVGSKALHETIATHQNVGNREKIEYR-REDDRFHTGYVTDSL 711
 QY 686 KSEVIVGSGQNFVFKSGKFNDIFHSGEGLDLDGAGDDRLFGCKGNDRLSGEGDILL 745
 DB 712 KSVIEIIGSQFNDFIFGSGQFDFVFGNGVDTIDNGDDHLLFGGAGDDVDIDGNGNPL 771
 QY 746 DGGSGDDVLNGAGNDVYIFRKGDCNDTLDYGTGNDKLAFAFADANISDTMIERTKEGIIVK 805
 DB 772 VGGTGNDIISGGKNDIYVHTGDNDSITDSGGQDKLAFSDVNLKDLTFKKVDSLEI- 830
 QY 806 RNDHSGSINIPRWY-----ITSNLQYQSNKTDHKEIQLEIGKDGSYITSQIDKILQDKD 861
 DB 831 INQKGVRIQNWFEDEDLASTVANYKAT-NDRKIEEIIKGGERITSEQVDKLI--KEG 887
 QY 882 GTVITSQELKKLADENKSKQSASDIASSLNKLVCMSMALFGTANSVSN 910
 DB 888 NNQISAEALSQVNDYNTSK-DRQNVNSLAKLITSSVGSFTSSSDFRNN 935

RESULT 32

ID AAW22159 standard; Protein; 1049 AA.

AC AAW22159;

DT 16-FEB-1998 (first entry)

DE ApxIIB protein.

KW RTX toxin; apxICA gene; apxIIB gene; apxIIB/C gene; apxIIBABD gene;
 KW repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
 KW therapy; A. pleuropneumoniae infection; swine pleuropneumonia.

XX Actinobacillus pleuropneumoniae.

XX CA2170839-A.

XX 02-SEP-1996.

XX 01-MAR-1996; 96CA-2170839.

XX 01-MAR-1995; 95US-0396244.

XX (UYGU-) UNIV GUELPH.

XX MacInnes J, Mallard B, Ricciatti P, Rosendal S;

XX WPI; 1997-245536/23.

XX N-PSDB; AAT73220.

PT Preparations of microorganisms producing cell-associated RTX toxins
 PT - especially for production of vaccines against swine
 PT pleuro-pneumonia

PS Disclosure; Pages 107-110; 151pp; English.

CC AAW22151-W22161 represent A. pleuropneumoniae RTX (repeat in toxins)
 CC toxins. These sequence are encoded by the apxICA, apxIIB, apxIIBABD, and
 CC apxIIBABD genes (see AAT73217-T73220), and can be expressed by
 CC microorganisms used in the preparations of the invention. The
 CC preparations are bacterial preparations comprising one or more isolated
 CC and purified strains of a microorganism that produces one or more RTX

CC toxins, where the strains have at least one cell-associated RTX toxin.
 CC The preparations are used for production of vaccines for the prophylaxis
 CC and treatment of infectious diseases caused by microorganisms that
 CC produce RTX toxins, where the strains have been attenuated or
 CC inactivated. The vaccines are preferably against Actinobacillus
 CC pleuropneumoniae infection (swine pleuropneumonia). It has been found
 CC that A. pleuropneumonia produces significant quantities of
 CC cell-associated RTX toxins when cultured under certain conditions, and
 CC that the whole-cell protein composition of the cultures corresponds to
 CC the whole-cell protein profiles obtained from cells recovered at
 CC necropsy from the pleural fluid of infected swine. Vaccination with a
 CC bacterin prepared from heat-inactivated cultures having significant
 CC quantities of cell-associated RTX toxins gives significant protection of
 CC swine against challenge with homologous strains.

XX Sequence 1049 AA;

Query Match 41.4%; Score 1924; DB 18; Length 1049;
 Best Local Similarity 43.0%; Pred. No. 2.5e-111;
 Matches 428; Conservative 152; Mismatches 296; Indels 120; Gaps 21;

QY 8 KSNTOAGLNSTKSLK-----NLYLAIPKDYDPKQGTGLNDFI 45
 DB 19 KRQVKKGVDVTKNGLOYGVSOAKLQALAAAGKAVQKGNKLVLPVPEYDGSVGNGFDFLV 78
 QY 46 KADELGIARLAEPEPHETAKKSVDTVNOFLSLTQTGIALSATKLEKLOKHSTNKLA 105
 DB 79 KAABELGIQVKNVRNELEVAHKSIGTADOFPLGUTERGLTLFAPQLDQFLQKHGISNVV 138
 QY 106 GLDSVENIDRKLKASNVLSLSSFLGTALAGIELDSLIRKKGDAAPALAKASIDLIN 165
 DB 139 GSSTGDVAVS-KLAKSQTIIISGIVGLTVLAGINLNEAIIISGSELE-LAEAGVSLASEL 196
 QY 166 IGNUSTQTTIEAPSSQAKLGSTISOAKGFSNTGNKQNLN---FKTNLGLBIIITGL 222
 DB 197 VSNIAKGTITTDFTTQIQNFKLAENAKGLGGVGRQLQNTISGALSKTGLGLDIISL 256
 QY 223 SGISAGFALADKNASTCKKVAAGFELSNOVIGNVTKAISVYLAORVAAGLSTTGAAAL 282
 DB 257 SGVTRSFALRNKNASTSTKVAAGFELSNOVIGGTTKAVSSYILAORUAGLSTTGPAAL 316
 QY 283 ITSSIMLAISPLAFMNAADKFNHANALDEFAPKFRKFGYDGDHLLAEYQGVGTIEASLT 342
 DB 317 IASSISLAISPLAFRLVADNFNRSKEIGFAERFKLGYDGDKLLSEFYHEAGTIDASIT 376
 QY 343 TISTALGAVSAGVSAAGVAVGTPPIALLVAGVTGLISGLEASKQAMFESVANRLOOKI 402
 DB 377 TISTALSAIAAGTAAASAGALVGAPITLLVTGITGLISGLEFSGKQPMLDHVASKIGNKI 436
 QY 403 LEWEKQNGQNYFDKGYDSRYAAVLANNLKFLSELNKELEAEVIAITQORWNNNIGELA 462
 DB 437 DEWEKKY-GKNYFENGYDARHKAFLDSFSLSSFNKQYETERAVLITQQRWDEYIGELA 495
 QY 463 GITKLERIKSGKAYADAFEDGKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFT 518
 DB 496 GITCKGDKLSGKAYVDYFQEGKLEKPPDDFSKVVPDPTKGEIDISNS--QTSLLAKFV 553
 QY 519 SPLLTAGTESRRLTNGKYSYINKLKFRGVKNQVQTDG--EASSKLDPFSKVQR----- 570
 DB 554 TPLLTPGTESRRTQTGKYEYITKLVVVKGDKW-VVNGVKDKGAVDYTNLIQHAHSSS 612
 QY 571 VAETEGTDEGLIYNKAGNDDIFVGQCKNIDGSDGHDHDFVFSKDGFGNITVDGTSAT 630
 DB 613 VARGEYREVLVSHLGNLGNKDFVLAAGSAEIHAGEHDVVVYDKT-DTGLVLVDGTQAT 671
 QY 631 EAGSYTVNRKVARG-DIYHEVVRQETKVGKRTETIQVRDYELRKVG-YGYQSTDNLSKV 688
 DB 672 EQGRYSVTRELSGATKILREVIKNQYAVGKREETLEYRDVELATQSGNSNLKAHDELHSV 731
 QY 689 BEVTGSGFNDVFKGSKFNDIFHSGEGLDLDGAGDDRLFGCKGNDRLSGPD----- 739
 DB 732 EE-IGSNQRDFKSGKFRDIPFHGADGDDLLNGNDGDDILYDGKNGDELGRGNDGNDQLYGG 790

Qy 740 EGD-----DLDGSGDDV 753
 Db 791 EGDGKLGGNNYLSGDDGNDLQVNGFVLRGKGDDKLYSGSGDLDGEGENDY 850
 Qy 754 LGGAGNDVYIFRKGNDNTLYD---GTGNDKLAADANISDIIMERTKEGIIVK--RND 808
 Db 851 LEGGDSDFYVYRSTSGNHTIYDQKASDSDKLYSLDLSFDNLVKRVNDLPRSNNS 910
 Qy 809 HSGSINPRWITSNLQYQSNKTDHKEIQILGKGSYITSDQIDKLODKDGTVITSQ 868
 Db 911 NSGVLTITKDFKGNYSN-----HKIEQIVDKNGRKLTAGNLGNPHD-----TOQASS 959
 Qy 869 ELKKLADENKSQLSASDIASSLNKLVSMAFGTA 904
 Db 960 LKKNVTQESNLS--LKTELKIIITNAGNFGVA 993

RESULT 33
 AAY51412
 ID AAY51412 standard; protein; 1049 AA.
 XX AAY51412;
 AC
 XX
 DT 05-MAY-2000 (first entry)
 XX
 DE A. pleuropneumoniae apxIIIA protein.
 KW RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
 KW antiinflammatory; antiarthritic; antiabortive; treatment; pneumonia;
 KW pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
 KW shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
 KW urinary infection; peritonitis; meningitis; gastroenteritis;
 KW passive immunization; apxIIIA.
 XX
 OS Actinobacillus pleuropneumoniae.
 XX
 XX US6019984-A.
 XX
 XX 01-FEB-2000.
 XX
 XX 23-DEC-1996; 96US-0772270.
 XX
 XX 01-MAR-1995; 95US-0396244.
 XX
 XX (UTGU-) UNIV GUELPH.
 XX
 XX Mallard B, Rosendal S, MacInnes J, Ricciatti P;
 XX
 XX WPI: 2000-146864/13.
 XX
 XX N-PSDB; AAZ88587.
 XX
 XX Bacterial preparation comprising microorganisms which produce a member
 PT of the Repeats in Toxins (RTX) family, useful for treating swine
 PT pleuropneumonia, arthritis in swine, shipping fever and abortion in
 PT cattle, and sleepy foal disease -
 XX
 XX Disclosure; Column 77-84; 96pp; English.

This invention describes a novel bacterial preparation (I) which
 comprises one or more isolated and purified strain(s) of a microorganism,
 cultured in tryptone yeast extract (TYE) broth, which produces one or
 more RTX toxins (belonging to the family of toxins referred to as Repeats
 in Toxins), where the strain(s) have at least one RTX toxin which is
 cell-associated. The products of the invention have immunostimulatory,
 antimicrobial, antiinflammatory, antiarthritic and abortifacient activity.
 The bacterial preparation may be used as vaccines for the prophylaxis and
 treatment of infectious diseases caused by strains of microorganisms
 which produce one or more RTX toxins. The infectious diseases are swine
 pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
 septicemia, nephritis, endocarditis and arthritis in piglets; shipping
 fever and abortion in cattle; whooping cough, sleepy foal disease or
 joint ill (purulent nephritis, arthritis) in foals; septicemia,
 polyarthritis and abortion in horses; and urinary infections,

CC peritonitis, meningitis, and gastroenteritis. The bacterial preparations
 CC may also be used to prepare antibodies which may be used as a means of
 CC passive immunization. This sequence represents the Actinobacillus
 CC pleuropneumoniae apxIIIA protein described in the method of the
 CC invention.
 XX
 SQ Sequence 1049 AA;
 Query Match 41.4%; Score 1924; DB 21; Length 1049;
 Best Local Similarity 43.0%; Pred. No. 2.5e-111; Indels 120; Gaps 21;
 Matches 428; Conservative 152; Mismatches 296;
 Qy 8 KSNIQAGLNSTKSGLK-----NLYLAIPKDYDPKQGTLDNDFI 45
 Db 19 KRQVKGYDVTNKLQYGVSOAKLQALAAKAVQKYGNKLVLI PREYDGSVNGFFDLV 78
 Qy 46 KAADELGIARLAEPNHTETAKSVDTVNOFLSLTOTGTIAISATKLEKFLQKISTNKLAK 105
 Db 79 KAAEELGIQVYVNRNELEVAHKSLGTAQDFLGTERTGLTFLAPQLDQFLQKHSKISNVV 138
 Qy 106 GLDSVENIDRKLKGSANVLSLSSFLGTALAGTFLDSLKKGDAAPDALAKASIDLINEL 165
 Db 139 GSSTGDAVS-KIAKSTIISGIIOSVLTGVLGINLNEAIIISGSELE-LAEAGVSLASEL 196
 Qy 166 IGLNSOSTOTIEAFSSQAKLGSTISOAKGFSNIGNKLNLN---FSKTNLGLIEITGLL 222
 Db 197 VSNIAKGTITIDFTQIQNFGKLAENAKLGGVGRQLQNISSALSCTGLGDIISLL 256
 Qy 223 SGISAGFALADKNASTGKVAAGFELSNQVGNVTKVAISSYVLAQVAAVAGSLTGAVAL 282
 Db 257 SGVTRSFALRNKNASTSTKVAAGFELSNQVIGITKAVSSYILAQRRLAGSLTGTGAAL 316
 Qy 283 ITSSIMLATSPLAPMNAADKFNANALDEFKQFRKFGYDGDHLLAERYQYRGVGTIEASLT 342
 Db 317 IASSISLATSPLAPLRVADNFRNKRKEIGEPAERFKLGYDGDKLLSEFYHEAGTIDASIT 376
 Qy 343 TISTALGAVSAGVSAAGVAVGTPIALLVAGVTGLISGLEASQAMFESVANRLQGI 402
 Db 377 TISTALSAIAAGTAASAGALVGPITLLVTGITGLISGLEFSKQPMLDHVASKIGNKI 436
 Qy 403 LEWEKQNGQNYDFKGYDSRYAAYLANNKLFSELNKELEAEARVIAITQORWNNIGELA 462
 Db 437 DEWEKKY-GKNYFENGYPDARHKAFLDSFSLSSFNKQYETERAVLITQORWDEYIGELA 495
 Qy 463 GITKLGERIKSGKAYADAFEDGKKVEAG-----SNITLDAKTGIIIDISNSNGKQTALHET 518
 Db 496 GITKGDKLSSGKAYVDYFOEGKLLKPKDPDFKVVDPFKGEIDISNS--QTSTLLKPV 553
 Qy 519 SPLLTAGTESRERLTNGKSYINKLFRVKNWQVTDG--EASSKLDFFSKVIQR----- 570
 Db 554 TPLLTGCTESRERTQTKYEYITKL VVGKDKW-VVNGVKDKGAVDYTNLIQHAHSSS 612
 Qy 571 VAETEGTDEGLIVNAKAGNDIIFVGGKKNIDGGDGHDRVYFYSKGGFGNITVDGTSAT 630
 Db 613 VARGEERYEVRLVSHLGNKNDKVFALAAGSAEIAHAGEGHVYVYDKT-DTGLLVIDGTGKAT 671
 Qy 631 BASGYTVNRKARG-DIYHEWVKRQETKVKRYETIOYRDYELERKVG-YGVQSTDNLSKV 688
 Db 672 EQGRYSVTRELSGATKLREVIKNQYAVGRBETLEIRDYELTQSGNSNLKAHDELHVS 731
 Qy 689 EEVIGSQFNDVFKSGFNDFIFHSGEGDDLLDGGAGDRLFGKGNDRLSGD----- 739
 Db 732 EE-IGSNORDEFKSGKFRDIFHGADGDDLLNGNDGDDLLYGDKNDELGRGNDQLYGG 790
 Qy 740 EGD-----DLDGSGDDV 753
 Db 791 EGDDKLLGGNNYLSGDDGNDLQVNGFVLRGKGDDKLYSGSGDLDGEGENDY 850
 Qy 754 LGGAGNDVYIFRKGNDNTLYD---GTGNDKLAADANISDIIMERTKEGIIVK--RND 808
 Db 851 LEGGDSDFYVYRSTSGNHTIYDQKASDSDKLYSLDLSFDNLVKRVNDLPRSNNS 910
 Qy 809 HSGSINPRWITSNLQYQSNKTDHKEIQILGKGSYITSDQIDKLODKDGTVITSQ 868

Db 911 NSGVLTIDKDFKGSYN-----HKIEQIVDKNGKRLTAGNLGNPHD-----TQOASS 959
Qy 869 ELKKLADENKQKLSASDIAGSLNKLKVGSMALFGTA 904
Db 960 LLKNVTQEQNESNLSS--LKTELKGIITNAGNFGVA 993

RESULT 34
ID AAR54781 standard; Protein; 1244 AA.
XX AAR54781;
XX AC
XX DT 25-MAR-2003 (updated)
XX DT 15-OCT-1994 (first entry)
XX Leukotoxin AppIIIA.
XX Leukotoxin; AppIIIA; pleuropneumonia; vaccine; diagnostic.
XX OS Actinobacillus pleuropneumoniae.
XX FH Key Location/Qualifiers
XX FT Domain 327..345
XX FT /label= transmembrane domain
XX FT Domain 488..503
XX FT /label= transmembrane domain
XX FT Domain 570..587
XX FT /label= transmembrane domain
XX WO9409821-A1.
XX PD 11-MAY-1994.
XX PF 04-NOV-1993; 93WO-US10500.
XX PR 05-NOV-1992; 92US-0972229.
XX PR 03-JUN-1993; 93US-0072285.
XX (CORR) CORNELL RES FOUND INC.
XX Chang Y;
XX WPI; 1994-167130/20.
XX N-PSDB; AAQ64827.
XX DNA encoding Actinobacillus pleuropneumoniae leukotoxin - used to
XX prepare prods. for use in vaccines for porcine pleuropneumonia
XX and as diagnostic reagents
XX Disclosure; Fig 3; 65pp; English.
XX The AppIIIA gene having the sequence given in AAQ64827 was isolated
XX from phage lambda clones yfc 26-28 and yfc 31-32 of a genomic
XX library of A. pleuropneumoniae ser. 2. The gene encodes a novel
XX leukotoxin (sequence AAR54781).
XX (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 1244 AA;

Query Match 41.4%; Score 1924; DB 15; Length 1244;
Best Local Similarity 43.0%; Pred. No. 3.1e-111;
Matches 428; Conservative 152; Mismatches 296; Indels 120; Gaps 21;
Qy 8 KSNIOAGLNSTKSGLK-----NLVLAIPKDYDPQKGTLDNFI 45
Db 192 KRQVKGYDVTNGLOYSQAKLQALAAAGKAVQKYNKLVLPKEYDGSVNGGFFDLV 251
Qy 46 KAADELGIARLAEPNHTETAKSVYDVTNQFSLTQTGTGIAISATKLEKPLQKHSNKLAK 105
Db 252 KAAEELGIQVYVNRNELEVAHKSGLTADQFLGLTERGLTLPAPQLDQFLQKHSISNW 311

Qy 106 GLDSVENIDRKLGKASNVLTSLSSFLGTALAGIELDSLIIKKGAAPALAKASIDLINEL 165
Db 312 GSSTGDAVS-KLAKSQTIIISGIVSLGTVLGINLNEALISGSELE-LAAGVSLASEL 369
Qy 166 IGNLSQSTOTIEAFSSQLAKLGSTISQAKFSNIGNKLONLN---FSKTNLGLBIIITGLL 222
Db 370 VSNIAKGTITIDAFITQIQNFGLAENAKGLGGVGRQLQNISSGALSALSKTGLGLDISSLL 429
Qy 223 SGISAGFALADKNASTGKVAAGFELSNOVIGNVTKAISSVLAORVAAGLSTTGVAAL 282
Db 430 SGVTRSFALRNKNASTSTKVAAGFELSNOVIGITKAVSSYILAORLRAGLSTTGPAAAL 489
Qy 283 ITSSIMLAISPLAPMNAADKFNHANALDEFAKQKFGYDGDHLLAYQORGVGTIEASLT 342
Db 490 IASSISLAISPLAFLRVADNFRNSKEIGEFAERFKLGYDGDKLLSEFFYEAGTIDASIT 549
Qy 343 TISTALGAVSAGVSAAGVAVGTPPIALLVAGVTGLISGILEASKQAMFESVANRLOCKI 402
Db 550 TISTALSAIAAGTAASAGALVGAPITLLVTGITGLISGILEFSKQPMLDHVASKIGNKI 609
Qy 403 LEWEKQNGQNYFDKGYDSRYAAVLANNLKFELSELNKELEAERVIAITQORWDNNIGELA 462
Db 610 DEWEKKY-GKNYFENGYDARHKAFLDSFLLSSFNKOYETERAVLITQORWDEYIGELA 668
Qy 463 GITKGERIKSGKAYADAFEDGKKVEAG---SNITLDAKTGIIDISNSNGKKTQALHFT 518
Db 669 GITGKDKLSSGKAYVDFQEGKLEKPPDDFSKVDFPTKGEIDISNS--QTSTLLKXFV 726
Qy 519 SPLTAGTRESRLTNGKYSYINKLFGRVKNQVOTDG--EASSKLPFSKVIQR----- 570
Db 727 TPLLTPTGESRERTQTKYEYITKLVVKGDKW-VVNGVKDKGAVDYDTNLIQAHISSS 785
Qy 571 VAETEGTDEGLIVNAKAGNDDIFVGQKKNIDGCGHDRVYFSKDGFGNITVDGTSAT 630
Db 786 VARGEERYEVLVSHLGNLGNKVFELAAGSABIHAGEGHVVYYDKT-DTGLLVIDGTGKAT 844
Qy 631 BAGSYTVNRKVARG-DIYHEVVKROETKVGKRTETIOVRDYELRKVG-YGYOSTDNLKS 688
Db 845 EGRYSVTRELSGATKILREVINKQYAVGKRETFEYRDYELTQSGNSNLKAHDELHSV 904
Qy 689 BEVIGSQFNDVFKSGKFNDIFHSGEGDDLDDGGAGDRLFGKGNDRLSGD----- 739
Db 905 EE-IGSNQRDEFKSGKFRDIFHGADGDDLNGNDGDDILYDGKGNDELGRDNGNDLYGG 963
Qy 740 EGD-----DILDCSGSDV 753
Db 964 EGGDKLLGNGNNYLSGGDGNDELQVLGNGFNLRGKGGDKLYGSSGSDLLDGGEGNDY 1023
Qy 754 LAGGAGNDVYIFRKGDGNDLYD---GTGNDKLAFAADANISDIERTKEGIIIVK--RND 808
Db 1024 LEGDGSDFVYRSTSGNHTIYDQKASDSKLYSLSPNLIIVKRVNDNLEFRSNNS 1083
Qy 809 HSGSINIPRWYITSNLQYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQ 868
Db 1084 NSGVLTIDKDFKGSYN-----HKIEQIVDKNGKRLTAGNLGNPHD-----TQOASS 1132
Qy 869 ELKKLADENKQKLSASDIAGSLNKLKVGSMALFGTA 904
Db 1133 LLKNVTQEQNESNLSS--LKTELKGIITNAGNFGVA 1166

RESULT 35
ID AAR22152 standard; Protein; 1022 AA.
XX AAR22152;
XX AC
XX DT 16-FEB-1998 (first entry)
XX XX
DE ApxIA protein.
KW RTX toxin; apxICA gene; apxIBD gene; apxIIAB'C gene; apxIIABCD gene;
KW repeat in toxins toxin; cell-associated RTX toxin; vaccine production;


```
QY 452 ORWNNIGELAGITKLBRIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNGKK 511
DB :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
125 QOWDNNIGDLAGISLGEKVSGLKAYVDAFEGKHAKADLVQSDANGIIDVNSGKAK 184
QY 512 TQALHTSPLITAGTESRLTNGKYSYINKLKFGVRKRWQVTDGEASKLDFSKVQRV 571
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
185 TQHILFRTPLLTPGTEHREVRQTKYEVITKLNINRVDSWKITDGAASSTFDLTNNVQRI 244
QY 572 -----RETEGTDEIGLIVNAKAGNDIIFVCGHNMNIDGGGHDHVRFYSGKGGFCNITV 624
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
245 GIELDAGNVTKETKIIAKLGEKDDNVFVSGGTETIDGEGYDRVHYR-GNYGALTI 303
QY 625 DGTSAEAGSYVNRKVARGDIIYBEVKRQETKVGKRTETIQYRDYELRKVGYGQSTDN 684
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
304 DATKETEGSYVNRVETGKALHEVTSHTALVGNREEKIEYR-HSNNQHAGYTYKDT 362
QY 685 LKSEEVIGSQPNDFVFGSKFNDFPHSGEGDDLLDGGAGDDRLFGKGNDRLSGDEGDDL 744
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
363 LKAVEEIIIGTSHNDIFKSGKFNDAFNGGDGVDTIDGNDGNRLFGKGDDILDGGNGDDF 422
QY 745 LDGSGDDVLNGGAGNDVYIFRKGDCNDTLVDGTGNDKLAADANISIMIERTEKGLIV 804
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
423 IDGKGNLHGGKGGDDIFVHRKGGNDIITSDGNDKLSFSDSNLKDITFEKVKNHLVI 482
QY 805 KENDHSGSINIPRWY----ITSNLQYOSNKTDKHIEOLIGKGSYITSDQIDKILQDK 860
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
483 -TNSKKEKVTQNWFEADFAKEVNYKATK-DEKIEHIIQNGERITSKQVDDLI--AK 538
QY 861 DGTVITTSQELKKLADENKSKLSADSIASSLNKLVGSMALFGTANSVSNALQIPTQ 920
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
539 GNGKITQBELSKVDVNYELLKHS-KNVTNSLDKLTSSVSAFTSSNDSNRNLVAPTMLDQ 597
QY 921 GI 922
DB :
598 SL 599

RESULT 39
AAR86998
ID AAR86998 standard; Protein; 758 AA.
AC AAR86998;
XX
DT 04-JUL-1996 (first entry)
DE Enterohaemorrhagic E.coli hlyA gene product.
KW Enterohaemorrhagic Escherichia coli; virulent; EHEC; O157:H7 serotype;
KW detection; probe; primer; hlyA gene; enterohaemorrhagic colitis;
KW haemolytic uremic syndrome; mesenteric adenitis.
XX
OS Escherichia coli (enterohaemorrhagic).
XX
PN US5475098-A.
XX
PD 12-DEC-1995.
XX
PF 14-JUN-1994; 94US-0258188.
XX
PR 14-JUN-1994; 94US-0258188.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Hall RH, Xu JG;
XX
XX WPI; 1996-048546/05.
DR N-PSDB; AAT08098.
XX
PT Enterohaemorrhagic E. coli (EHEC) nucleic acid sequences - useful
PT for probe and primer design for sensitive and specific detection of
PT EHEC
XX
PS Claim 1; Columns 37-42; 32pp; English.
```

```
XX Enterohaemorrhagic E.coli (EHEC) associated with enterohaemorrhagic
CC colitis, haemolytic uremic syndrome and mesenteric adenitis have
CC been found to carry a hlyA gene and a hlyB gene, separated by an
CC intergenic region. The hly genes and the intergenic region are
CC absent from bacteria not associated with these diseases and so
CC provide a useful target for detecting EHEC pathogens, esp. O157:H7
CC serotype E.coli. The present sequence is that of the protein
CC encoded by the EHEC hlyA gene.
XX
SQ Sequence 758 AA;
Query Match 29.6%; Score 1376.5; DB 17; Length 758;
Best Local Similarity 40.5%; Pred. No. 2.2e-77;
Matches 304; Conservative 129; Mismatches 225; Indels 93; Gaps 16;
QY 229 PALADKNASTGKVKVAGFELSNOVGNVTKAISVLAORVAAGLSTTGVAALITSSIM 288
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2 FILGNSDAHTGTAAAGIELTTVLGNVGKAVSOYILAQRMAQGLSTTAASAGLITSAMV 61
QY 289 LAISPLAFMAADKFNHANALDEPAKOFREGYDGDHLLAEYQRGVGTIEASLTITSTAL 348
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
62 LAISPLSFLAAADKFERAKQLESYSEFKLNVEGDALLAFAFKBTGAIDAAALTINTVL 121
QY 349 GAVSAGYSAAGVSAVGTPIALIVAGVTGLISGLEASKQAMFESVANRLQCKILEWEKQ 408
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
122 SSVSAGYSAASSALIGAPISMLVSALTGTISGLEASKQAMFEHVAEKFAARINEWEKE 181
QY 409 NGGQNYFDKGYDSRYAAYLANNKFLSELNKELEBAERVIAITQORWNNNIGELAGITKLG 468
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
182 H-KNYPFENGYPDARHAAFLDLSLADFSRQHAVERAVAITOOHMDKIELAGITRNA 240
QY 469 ERIKSGKAYADAFEDGKKVEAG---SNIITLDAKTGIIDISNSNGKKTQALHFTSPLTA 524
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
241 DRSQSGKAYINYLENGGLLEAQPKFTQQVDFPDKGTIDL--STGNVSSVTFITPTFT 298
QY 525 GTSERRLTNGKYSYINKLKFRVKNQVTDGEASSK--LDFSKVIQRAVTEGTDEIGL 582
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
299 GEEVREKQSGKYEYMTSLIVNGKDTWSV-KGIKNHKGVDYDYSKLIQFVEKNNKHQAARI 357
QY 583 IVNAKAGNDIIFVCGHNMNIDGGGHDHVRFYSGKGGFCNITVDGTSAGSYTVNRKVA 642
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
358 ISELGDKDDVYVSGAGSEVFAGEGYDVTSYNKT-DVGKLTIDATGASKPGEYIVSKNM- 415
QY 643 RGD--IVHEVVKROETKVGKRTETIQYRDYELRKVGYGYSTDLKSVVEVIGSQFNDVF 700
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
416 YGDVKVQLQEVVKEQSVGKRTEKIQYRDFEFTFGGIPYDVIDNLHLSVEELIGKHDEF 475
QY 701 KGSKFNDFHSG-----EGDDLLDGGAGDDRLFGKGNDRLSGDEGD 742
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
476 KGGKFNDI FHGAGNDYIEGNYGNDRLYGDDGDDYISGGQGDQLFGSGNDKLSGDDGN 535
QY 743 -----DILDGGSGDDVLNGAGNDVYIFR 766
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
536 NYLTGSGNDELQAHGAYNIIISGTTGDDKLYGGGGIDLLDGGEGNDYILNGFGNDIYVYG 595
QY 767 KGDGNDTLVD-GTGNDKLAFADANISIDMIERTKEGIIVKR-----NDHSSGIN 814
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
596 QNYGHHTIADGEGKGRHLHSDISFDDIAFRVGNNDLIMKNAINGVLSFNESNDVNG-IT 654
QY 815 IPRWYITSNLQYOSNKTDKHIEOLIGKGSYITSDQIDKILQDKGDTVITTSQELKKLA 874
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
655 PKWFAKD-----ASGADNHLVEVITDKDGREI---KVDKIPHNNE-----RSGVIKASN 702
QY 875 DENKSQKLSASDIASSLNKLVGSMALFGTAN 905
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
703 IASEKNMVNITSVANDINKIISVSGFDSGD 733

RESULT 40
AAE04637
ID AAE04637 standard; Protein; 450 AA.
XX
```

AC AAE04637;
XX
XX 10-SEP-2001 (first entry)
XX
XX Pasteurella haemolytica modified leukotoxin 50 (lkt50) protein.
XX
XX Leukotoxin 50; lkt50; respiratory disease; infection; therapy;
XX immunostimulant; antibacterial; vaccine; transgenic plant;
XX animal feed; mutant; muten.
XX
XX Pasteurella haemolytica.
XX Synthetic.
XX
XX WO200144289-A2.
XX
XX 21-JUN-2001.
XX
XX 15-DEC-2000; 2000WO-CA01498.
XX
XX 17-DEC-1999; 99US-0172148.
XX
XX (UYGU-) UNIV GUELPH OFFICE.
XX
XX Lo RYC, Shewen PE, Lee RMH, Hodgins D, Strommer JN;
XX
XX WPI; 2001-408470/43.
XX N-PSDB; AAD08976.
XX
XX Modified leukotoxin polypeptide is useful in a vaccine to prevent or
XX treat Mannheimia (Pasteurella) infection (particularly M. haemolytica
XX infection), and disease associated with a leukotoxin, e.g., respiratory
XX disease
XX
XX Claim 4; Fig 12; 70pp; English.
XX
XX The present sequence is Pasteurella (Mannheimia) haemolytica
XX modified leukotoxin-50 (lkt50) protein. The modification comprises
XX the removal of amino acids within the hydrophobic transmembrane
XX domain of a full length leukotoxin protein. Modified leukotoxin
XX sequences are used in vaccines to treat or prevent diseases associated
XX with leukotoxin, e.g., respiratory disease, and Mannheimia infection
XX (particularly M. haemolytica infection). In addition, the vaccine
XX is used to prepare a medicament. Furthermore, the plant transformed
XX with modified leukotoxin sequences is fed to an animal such as a
XX ruminant, to prevent or treat respiratory diseases.
XX
XX Sequence 450 AA;

Search completed: February 17, 2004, 10:13:13
Job time : 55 secs

Query Match 25.0%; Score 1159.5; DB 22; Length 450;
Best Local Similarity 49.6%; Pred. No. 3.6e-64;
Matches 225; Conservative 76; Mismatches 136; Indels 17; Gaps 7;

Qy 433 FLSELNKELEAERVIAITQQRWNNIGELAGITKLGERIKSKAYADAFEDGKKEAGSN 492
Db 1 FLLNLNKELEAERVIAITQQRWNNIGELAGITKLGERIKSKAYADAFEDGKKEAGSN 492

Qy 493 ITLDAKTIIDISNGKKTQALLFTSPLLTAGTESRRLTNGKYSINKLKFGRVKNWQ 552
Db 61 VOLDASANGIIDVSNKGAQTQHLFTPELLTPGTEHREVRVQTKYETITKLININRVDSWK 120

Qy 553 VTGGEASSKLFQSVIRV-----AETEGDEIGLIWNAKAGNDDIFVGQKKNIDCG 605
Db 121 ITDGNASTFLTNVQRIEGLDAGNVTYKTKETIIAKLGEEDNVFVSGTTEIDCG 180

Qy 606 DGHDRVYFKDGGFGNITVDGTSATFAGSYTVNRKVARGDIYHEVVRQETKVGKRTETI 665
Db 181 EGYDRVHYSR-GNYGALITDATKETEQGSYTVNRVETGKALHEVTSHTALVGNREKI 239

Qy 666 QYRDYELRVKGYGYOSTNLKSVREVIGSQFNDVFKGSKFNDIFHSGEGDILLDGGAGDD 725
Db 240 EYR-HSNQHGAGYVTKOTLKAVESIIIGTSHNDIFPKSKFNDAFNGGSDVDTIDGNDG 298

Qy 726 RLFGKGNDRLSGDGGDLLDGGSDVDLNGGAGNDVYIFRKGNDNTLYDGTGNDKLAF 785

THIS PAGE BLANK (USPTO)

Search notes

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:13:20 ; Search time 47 Seconds
(without alignments)
3130.628 Million cell updates/sec

Title: US-10-069-799-5
Perfect score: 927
Sequence: 1 MSNINVIKSNIOAGLNSTKS.....SSNALQIPTPTQGTILAPSV 927

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 927 | 100.0 | 927 | 22 | M. bovis Dalton 2d |
| 2 | 14 | 1.5 | 758 | 17 | Enterohaemorrhagic |
| 3 | 14 | 1.5 | 1049 | 18 | ApIIIB protein. |
| 4 | 14 | 1.5 | 1049 | 21 | A. pleuropneumonia |
| 5 | 14 | 1.5 | 1244 | 15 | Leukotoxin ApIIIA |
| 6 | 11 | 1.2 | 956 | 12 | APPA haemolysin an |
| 7 | 11 | 1.2 | 956 | 18 | ApIIIC protein. A |
| 8 | 11 | 1.2 | 956 | 21 | A. pleuropneumonia |
| 9 | 11 | 1.2 | 1022 | 18 | ApIIA protein. AC |

| | | | | | | |
|----|----|-----|------|----|----------|---------------------|
| 10 | 11 | 1.2 | 1022 | 21 | AAV51406 | A. pleuropneumonia |
| 11 | 11 | 1.2 | 1023 | 15 | AAV76991 | LhaA (low homology |
| 12 | 10 | 1.1 | 450 | 20 | AAE04637 | Pasteurella haemol |
| 13 | 10 | 1.1 | 490 | 20 | AAV33929 | Leukotoxin carrier |
| 14 | 10 | 1.1 | 544 | 17 | AAW03943 | LKT-GNRH chimeric |
| 15 | 10 | 1.1 | 544 | 19 | AAW79570 | LKT-GNRH chimeric |
| 16 | 10 | 1.1 | 608 | 22 | AAE04636 | Pasteurella haemol |
| 17 | 10 | 1.1 | 695 | 19 | AAW79573 | LKT-GNRH chimeric |
| 18 | 10 | 1.1 | 695 | 19 | AAV58361 | Leukotoxin/gonadot |
| 19 | 10 | 1.1 | 695 | 21 | AAV58361 | Leukotoxin/gonadot |
| 20 | 10 | 1.1 | 802 | 22 | ABB67818 | Gonadotropin relea |
| 21 | 10 | 1.1 | 924 | 12 | AAV10889 | Drosophila melanog |
| 22 | 10 | 1.1 | 924 | 14 | AAV42385 | Leukotoxin 352 enc |
| 23 | 10 | 1.1 | 924 | 14 | AAV42380 | Recombinant leukot |
| 24 | 10 | 1.1 | 924 | 14 | AAV42378 | Recombinant leukot |
| 25 | 10 | 1.1 | 926 | 12 | AAV14482 | LKT352. Pasteurel |
| 26 | 10 | 1.1 | 926 | 14 | AAV34545 | Leukotoxin 352 pro |
| 27 | 10 | 1.1 | 926 | 15 | AAV50291 | Recombinant leukot |
| 28 | 10 | 1.1 | 926 | 17 | AAV03945 | P. haemolytica tru |
| 29 | 10 | 1.1 | 926 | 19 | AAV79568 | Leukotoxin 352 pol |
| 30 | 10 | 1.1 | 934 | 17 | AAV07637 | P. suis leukotoxin |
| 31 | 10 | 1.1 | 936 | 14 | AAV34547 | GNRH-leukotoxin ge |
| 32 | 10 | 1.1 | 943 | 14 | AAV34546 | Somatostatin-leuko |
| 33 | 10 | 1.1 | 951 | 14 | AAV34548 | Rotavirus VP4-leuk |
| 34 | 10 | 1.1 | 953 | 11 | AAV07167 | 105kd PTX protein |
| 35 | 10 | 1.1 | 953 | 12 | AAV15159 | Leukotoxin from P. |
| 36 | 10 | 1.1 | 953 | 14 | AAV34865 | Leukotoxin protein |
| 37 | 10 | 1.1 | 953 | 15 | AAV60072 | PtXA protein of Pa |
| 38 | 10 | 1.1 | 953 | 22 | AAV04638 | Pasteurella haemol |
| 39 | 10 | 1.1 | 977 | 17 | AAV03942 | LKT-GNRH protein f |
| 40 | 10 | 1.1 | 977 | 19 | AAV79569 | LKT-GNRH chimeric |
| 41 | 10 | 1.1 | 1069 | 15 | AAV52748 | Bovine IFNGamma/LK |
| 42 | 10 | 1.1 | 1069 | 18 | AAV13867 | Chimeric protein # |
| 43 | 10 | 1.1 | 1069 | 21 | AAV21074 | Bovine gamma-IFN/P |
| 44 | 10 | 1.1 | 1098 | 13 | AAV22103 | Bovine IL-2 - LKT |
| 45 | 10 | 1.1 | 1098 | 15 | AAV52747 | Chimeric protein ch |
| 46 | 10 | 1.1 | 1098 | 18 | AAV13866 | Bovine IL-2/Pasteu |
| 47 | 10 | 1.1 | 1098 | 21 | AAV21073 | Bovine IL-2/Pasteu |
| 48 | 10 | 1.1 | 1334 | 12 | AAV14481 | LkA::lacZ fusion p |
| 49 | 10 | 1.1 | 1334 | 15 | AAV50290 | LkA::lacZ fusion |
| 50 | 10 | 1.1 | 1403 | 12 | AAV10890 | LkA::lacZ fusion |
| 51 | 10 | 1.1 | 1403 | 15 | AAV54629 | Mannuronan C-5-epi |
| 52 | 10 | 1.1 | 1403 | 19 | AAV79574 | Leukotoxin 101. S |
| 53 | 8 | 0.9 | 73 | 22 | AAV86900 | Human immune/haema |
| 54 | 8 | 0.9 | 97 | 23 | AAV07319 | Human ORFX protein |
| 55 | 8 | 0.9 | 186 | 21 | AAV75008 | Neisseria gonorrhe |
| 56 | 8 | 0.9 | 190 | 22 | AAV67626 | Propionibacterium |
| 57 | 8 | 0.9 | 251 | 21 | AAV75011 | Neisseria gonorrhe |
| 58 | 8 | 0.9 | 251 | 24 | AAV77888 | N. gonorrhoeae ani |
| 59 | 8 | 0.9 | 262 | 18 | AAV21779 | Protein encoded by |
| 60 | 8 | 0.9 | 341 | 23 | AAV66327 | Bifidobacterium lo |
| 61 | 8 | 0.9 | 368 | 23 | AAV49818 | Listeria monocytog |
| 62 | 8 | 0.9 | 480 | 23 | AAV93674 | Herbicideally activ |
| 63 | 8 | 0.9 | 553 | 15 | AAV54628 | Mannuronan C-5-epi |
| 64 | 8 | 0.9 | 580 | 23 | AAV38824 | A. niger tripeptid |
| 65 | 8 | 0.9 | 766 | 22 | AAV82742 | S. epidermidis ope |
| 66 | 8 | 0.9 | 805 | 23 | AAV39976 | Staphylococcus epi |
| 67 | 8 | 0.9 | 825 | 22 | AAV59301 | Drosophila melanog |
| 68 | 8 | 0.9 | 872 | 15 | AAV54631 | Fragment of mannur |
| 69 | 8 | 0.9 | 997 | 15 | AAV54630 | Mannuronan C-5-epi |
| 70 | 8 | 0.9 | 1262 | 18 | AAV13505 | B. bronchiseptica |
| 71 | 8 | 0.9 | 1644 | 18 | AAV13504 | B. bronchiseptica |
| 72 | 8 | 0.9 | 1644 | 18 | AAV13504 | B. bronchiseptica |
| 73 | 7 | 0.8 | 12 | 12 | AAV13283 | Parvovirus B19 epi |
| 74 | 7 | 0.8 | 25 | 15 | AAV53953 | Peptide reactive w |
| 75 | 7 | 0.8 | 34 | 23 | AAO21906 | BASB204 related pe |
| 76 | 7 | 0.8 | 34 | 23 | AAO21906 | BASB204 related pe |
| 77 | 7 | 0.8 | 35 | 23 | AAO21919 | BASB204 related pe |
| 78 | 7 | 0.8 | 38 | 23 | AAO21901 | BASB204 related pe |
| 79 | 7 | 0.8 | 38 | 23 | AAO21903 | BASB204 related pe |
| 80 | 7 | 0.8 | 38 | 23 | AAO21907 | BASB204 related pe |
| 81 | 7 | 0.8 | 38 | 23 | AAO21908 | BASB204 related pe |
| 82 | 7 | 0.8 | 38 | 23 | AAO21909 | BASB204 related pe |
| | | | 38 | 23 | AAO21910 | BASB204 related pe |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|----|----------|----------------------|-----|---|-----|-----|----|----------|--------------------|
| 229 | 7 | 0.8 | 284 | 24 | ABP57913 | X. campestris avrX | 302 | 7 | 0.8 | 407 | 23 | ABB54596 | Lactococcus lactis |
| 230 | 7 | 0.8 | 283 | 22 | AA394712 | Human protein sequ | 303 | 7 | 0.8 | 409 | 22 | ABG19403 | Novel human diagno |
| 231 | 7 | 0.8 | 283 | 24 | AAU70387 | Human adipocyte Se | 304 | 7 | 0.8 | 409 | 22 | AAG82454 | S. epidermidis ope |
| 232 | 7 | 0.8 | 286 | 22 | AAU39415 | Human polypeptide | 305 | 7 | 0.8 | 409 | 22 | AAG82606 | S. epidermidis ope |
| 233 | 7 | 0.8 | 286 | 22 | AA392555 | Human protein sequ | 306 | 7 | 0.8 | 415 | 23 | ABU51693 | Helicobacter pylor |
| 234 | 7 | 0.8 | 286 | 24 | ABP60400 | Pseudomonas fluore | 307 | 7 | 0.8 | 418 | 22 | ABG21503 | Novel human diagno |
| 235 | 7 | 0.8 | 300 | 21 | ACG14027 | Arabidopsis thalia | 308 | 7 | 0.8 | 424 | 22 | ABB61353 | Drosophila melanog |
| 236 | 7 | 0.8 | 301 | 23 | ABP27439 | Streptococcus poly | 309 | 7 | 0.8 | 424 | 23 | ABB07924 | Fruit fly pellino |
| 237 | 7 | 0.8 | 302 | 21 | ACG46712 | Arabidopsis thalia | 310 | 7 | 0.8 | 424 | 23 | ABP40164 | Staphylococcus epi |
| 238 | 7 | 0.8 | 302 | 22 | ACG82984 | S. epidermidis ope | 311 | 7 | 0.8 | 429 | 22 | ABG20995 | Novel human diagno |
| 239 | 7 | 0.8 | 305 | 21 | ACG10685 | Arabidopsis thalia | 312 | 7 | 0.8 | 435 | 23 | ABU11075 | Yeast selected int |
| 240 | 7 | 0.8 | 305 | 24 | AAE34574 | Human CD47 protein | 313 | 7 | 0.8 | 435 | 23 | ABG77204 | Selected interacti |
| 241 | 7 | 0.8 | 307 | 24 | ABU19021 | Pathogen specific | 314 | 7 | 0.8 | 439 | 22 | AA04085 | Botulism toxin C f |
| 242 | 7 | 0.8 | 308 | 22 | ACG92563 | C glutamicum prote | 315 | 7 | 0.8 | 439 | 23 | ABU50795 | Helicobacter pylor |
| 243 | 7 | 0.8 | 310 | 23 | ABR38856 | A. niger aspergill | 316 | 7 | 0.8 | 440 | 21 | AAU77135 | Synthetic botulinu |
| 244 | 7 | 0.8 | 311 | 21 | ACG29205 | Arabidopsis thalia | 317 | 7 | 0.8 | 440 | 22 | AA04091 | Botulism toxin hea |
| 245 | 7 | 0.8 | 311 | 22 | ABG79756 | Corynebacterium gl | 318 | 7 | 0.8 | 440 | 23 | ABU11271 | Yeast selected int |
| 246 | 7 | 0.8 | 311 | 23 | ABR77561 | Corynebacterium Cys | 319 | 7 | 0.8 | 440 | 23 | ABG77369 | Selected interacti |
| 247 | 7 | 0.8 | 314 | 21 | ACG08357 | Arabidopsis thalia | 320 | 7 | 0.8 | 448 | 23 | ABB54422 | Lactococcus lactis |
| 248 | 7 | 0.8 | 315 | 23 | ABP39676 | Staphylococcus epi | 321 | 7 | 0.8 | 451 | 22 | AAU01689 | Gene 31 human secr |
| 249 | 7 | 0.8 | 319 | 23 | ABR79408 | Mouse neuronal reg | 322 | 7 | 0.8 | 454 | 22 | ABB52906 | Escherichia coli p |
| 250 | 7 | 0.8 | 321 | 22 | AA065977 | Human secreted pro | 323 | 7 | 0.8 | 455 | 22 | AAU35965 | Helicobacter pylor |
| 251 | 7 | 0.8 | 321 | 22 | AA064717 | Gene 5 human secre | 324 | 7 | 0.8 | 458 | 22 | AAU35728 | Helicobacter pylor |
| 252 | 7 | 0.8 | 323 | 23 | ABP62901 | Human polypeptide | 325 | 7 | 0.8 | 458 | 22 | AA096223 | Putative P. abyssi |
| 253 | 7 | 0.8 | 323 | 23 | AAE17874 | Sequence of ompK36 | 326 | 7 | 0.8 | 460 | 22 | ABG17038 | Novel human diagno |
| 254 | 7 | 0.8 | 324 | 22 | AAU36286 | Pseudomonas aerugi | 327 | 7 | 0.8 | 460 | 22 | AAE05398 | Human interleukin- |
| 255 | 7 | 0.8 | 325 | 21 | ACG21863 | Arabidopsis thalia | 328 | 7 | 0.8 | 460 | 22 | AAE05401 | Human IRAK-4 mutan |
| 256 | 7 | 0.8 | 325 | 21 | ACG47535 | Arabidopsis thalia | 329 | 7 | 0.8 | 460 | 22 | AAE05403 | Human IRAK-4 polym |
| 257 | 7 | 0.8 | 325 | 23 | ABU50952 | Helicobacter pylor | 330 | 7 | 0.8 | 460 | 22 | AAE05404 | Human IRAK-4 polym |
| 258 | 7 | 0.8 | 326 | 21 | AA052500 | Helicobacter pylor | 331 | 7 | 0.8 | 460 | 23 | AAE24859 | Human IRAK4 protei |
| 259 | 7 | 0.8 | 328 | 21 | ACG14026 | Arabidopsis thalia | 332 | 7 | 0.8 | 460 | 23 | AAE24865 | Human IRAK4 protei |
| 260 | 7 | 0.8 | 329 | 20 | AA066356 | Partial human DNAX | 333 | 7 | 0.8 | 463 | 20 | AAU36773 | Chlamydia trachoma |
| 261 | 7 | 0.8 | 329 | 23 | AAE16097 | Human DNAX Toll li | 334 | 7 | 0.8 | 468 | 22 | AAU1895 | C. glutamicum meta |
| 262 | 7 | 0.8 | 331 | 22 | ABR71976 | Drosophila melanog | 335 | 7 | 0.8 | 468 | 22 | AAU79667 | Corynebacterium gl |
| 263 | 7 | 0.8 | 331 | 22 | ABG71976 | Novel human diagno | 336 | 7 | 0.8 | 468 | 22 | AAU79667 | Corynebacterium gl |
| 264 | 7 | 0.8 | 331 | 23 | AAU74332 | Human cytoskeleton | 337 | 7 | 0.8 | 471 | 18 | AAW16794 | Aspartase-470. Es |
| 265 | 7 | 0.8 | 331 | 23 | AAU74332 | Ramoplanin biosynt | 338 | 7 | 0.8 | 472 | 19 | AAW68394 | Clostridium botuli |
| 266 | 7 | 0.8 | 332 | 23 | AAU22146 | Drosophila melanog | 339 | 7 | 0.8 | 477 | 21 | AAU68317 | Escherichia coli B |
| 267 | 7 | 0.8 | 336 | 22 | ABR64015 | Arabidopsis thalia | 340 | 7 | 0.8 | 478 | 20 | AAU49079 | Escherichia coli a |
| 268 | 7 | 0.8 | 341 | 23 | ACG46711 | Lactococcus lactis | 341 | 7 | 0.8 | 478 | 20 | AAU33962 | E. coli aspartase |
| 269 | 7 | 0.8 | 343 | 20 | AAU56769 | C. trachomatis ser | 342 | 7 | 0.8 | 478 | 20 | AAU05731 | E. coli native app |
| 270 | 7 | 0.8 | 343 | 21 | ACG46710 | Arabidopsis thalia | 343 | 7 | 0.8 | 478 | 21 | AAU47163 | Arabidopsis thalia |
| 271 | 7 | 0.8 | 353 | 22 | ACG75406 | Human colon cancer | 344 | 7 | 0.8 | 478 | 21 | AAU79297 | Escherichia coli a |
| 272 | 7 | 0.8 | 354 | 23 | ABU51076 | Helicobacter pylor | 345 | 7 | 0.8 | 478 | 23 | ABU50702 | Helicobacter pylor |
| 273 | 7 | 0.8 | 358 | 23 | ABU05789 | M. tuberculosis an | 346 | 7 | 0.8 | 478 | 23 | ABB09551 | Escherichia coli K |
| 274 | 7 | 0.8 | 362 | 22 | AAU23384 | Novel human enzyme | 347 | 7 | 0.8 | 479 | 22 | ABG10527 | Novel human diagno |
| 275 | 7 | 0.8 | 365 | 21 | ACG10302 | Arabidopsis thalia | 348 | 7 | 0.8 | 487 | 22 | AA082711 | Aranicola proteoly |
| 276 | 7 | 0.8 | 366 | 23 | AAU72942 | Neisseria meningit | 349 | 7 | 0.8 | 492 | 24 | AAO16701 | Escherichia coli a |
| 277 | 7 | 0.8 | 366 | 24 | ABP77317 | N. gonorrhoeae ami | 350 | 7 | 0.8 | 499 | 21 | AA013363 | 1-deoxy-D-xyulose |
| 278 | 7 | 0.8 | 368 | 23 | ABP26740 | Streptococcus poly | 351 | 7 | 0.8 | 500 | 21 | AAU21939 | Arabidopsis thalia |
| 279 | 7 | 0.8 | 373 | 23 | ABF41917 | Human ovarian anti | 352 | 7 | 0.8 | 500 | 24 | AAE34903 | Mouse calcium aci |
| 280 | 7 | 0.8 | 379 | 22 | AAU18518 | Human cytoskeletal | 353 | 7 | 0.8 | 503 | 8 | AAU70219 | Sequence of serrap |
| 281 | 7 | 0.8 | 379 | 22 | AAU18543 | Human cytoskeletal | 354 | 7 | 0.8 | 504 | 22 | AAU14800 | Novel bone marrow |
| 282 | 7 | 0.8 | 380 | 20 | AAU06627 | C. elegans tyrosyl | 355 | 7 | 0.8 | 512 | 22 | ABB57797 | Drosophila melanog |
| 283 | 7 | 0.8 | 380 | 21 | AAU84308 | A tyrosylprotein s | 356 | 7 | 0.8 | 519 | 23 | AAE25289 | Human nucleic acid |
| 284 | 7 | 0.8 | 380 | 23 | ABR20551 | Cor diverged del | 357 | 7 | 0.8 | 519 | 23 | ABG66662 | Human fat responsi |
| 285 | 7 | 0.8 | 381 | 23 | ABR93951 | Herbicidally activ | 358 | 7 | 0.8 | 520 | 23 | ABG66661 | Rat fat responsive |
| 286 | 7 | 0.8 | 382 | 19 | AAW98312 | H. pylori GHPO 886 | 359 | 7 | 0.8 | 521 | 19 | AAU37934 | Mouse cAMP regulat |
| 287 | 7 | 0.8 | 382 | 22 | ABR63575 | Drosophila melanog | 360 | 7 | 0.8 | 523 | 22 | AA065800 | RP55-like protein |
| 288 | 7 | 0.8 | 382 | 22 | AAU41201 | Human polypeptide | 361 | 7 | 0.8 | 551 | 22 | AAU90877 | C glutamicum prote |
| 289 | 7 | 0.8 | 382 | 24 | ABU58164 | Human stress respon | 362 | 7 | 0.8 | 551 | 22 | AAU76764 | Corynebacterium gl |
| 290 | 7 | 0.8 | 383 | 21 | ACG10301 | Arabidopsis thalia | 363 | 7 | 0.8 | 557 | 19 | AAU71586 | Human hepatocyte n |
| 291 | 7 | 0.8 | 386 | 19 | AAW98237 | H. pylori GHPO 699 | 364 | 7 | 0.8 | 558 | 24 | ABP77082 | N. gonorrhoeae ami |
| 292 | 7 | 0.8 | 386 | 19 | AAW71489 | Helicobacter polyp | 365 | 7 | 0.8 | 560 | 15 | AA062523 | Hookworm anticoagu |
| 293 | 7 | 0.8 | 386 | 22 | AAU69494 | Human purified sec | 366 | 7 | 0.8 | 575 | 21 | AA021938 | Arabidopsis thalia |
| 294 | 7 | 0.8 | 387 | 23 | ABU51114 | Helicobacter pylor | 367 | 7 | 0.8 | 578 | 21 | AA052575 | Helicobacter pylor |
| 295 | 7 | 0.8 | 392 | 22 | ABR65237 | Drosophila melanog | 368 | 7 | 0.8 | 582 | 21 | AAU47162 | Arabidopsis thalia |
| 296 | 7 | 0.8 | 394 | 18 | AAW20797 | H. pylori transpor | 369 | 7 | 0.8 | 585 | 22 | AAU35539 | Haemophilus influe |
| 297 | 7 | 0.8 | 396 | 24 | ABP99303 | Orthomycin biosy | 370 | 7 | 0.8 | 587 | 22 | ABR52856 | Escherichia coli p |
| 298 | 7 | 0.8 | 396 | 24 | ABP99350 | Orthomycin biosy | 371 | 7 | 0.8 | 587 | 22 | ABR52856 | Escherichia coli p |
| 299 | 7 | 0.8 | 396 | 24 | ABP76716 | Streptomycetes virid | 372 | 7 | 0.8 | 589 | 22 | ABU52814 | Human metatolism-a |
| 300 | 7 | 0.8 | 400 | 21 | AAU82775 | Human chordin rela | 373 | 7 | 0.8 | 592 | 22 | AAU93825 | Human polypeptide, |
| 301 | 7 | 0.8 | 400 | 21 | AAU53033 | Human secreted pro | 374 | 7 | 0.8 | 594 | 24 | ABU06042 | N. meningitidis va |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|----|----------|---------------------|-----|---|-----|------|----|----------|----------------------|
| 375 | 7 | 0.8 | 598 | 21 | AAG21937 | Arabidopsis thalia | 448 | 7 | 0.8 | 969 | 22 | ABB58321 | Drosophila melanog |
| 376 | 7 | 0.8 | 609 | 19 | AAW71581 | Human hepatocyte n | 449 | 7 | 0.8 | 974 | 23 | AAU76874 | Human EphA full le |
| 377 | 7 | 0.8 | 617 | 18 | AAW27247 | Pseudomonas fluore | 450 | 7 | 0.8 | 982 | 22 | AB571714 | Drosophila melanog |
| 378 | 7 | 0.8 | 619 | 23 | ABU50976 | Helicobacter pylor | 451 | 7 | 0.8 | 986 | 22 | ABB57810 | Drosophila melanog |
| 379 | 7 | 0.8 | 619 | 23 | ABU52060 | Helicobacter pylor | 452 | 7 | 0.8 | 1000 | 22 | AAW47209 | Human NOV3 protein |
| 380 | 7 | 0.8 | 621 | 22 | ABB67394 | Drosophila melanog | 453 | 7 | 0.8 | 1001 | 23 | AAW68224 | Nilaparvata lugens |
| 381 | 7 | 0.8 | 622 | 22 | ABB60638 | Drosophila melanog | 454 | 7 | 0.8 | 1006 | 22 | ABB60312 | Drosophila melanog |
| 382 | 7 | 0.8 | 622 | 22 | ABB38995 | Human polypeptide | 455 | 7 | 0.8 | 1008 | 24 | ABB98843 | Human NEPHA. Homo |
| 383 | 7 | 0.8 | 627 | 22 | ABB63386 | Drosophila melanog | 456 | 7 | 0.8 | 1009 | 22 | AAU03553 | Human protein kina |
| 384 | 7 | 0.8 | 630 | 22 | ABB67395 | Drosophila melanog | 457 | 7 | 0.8 | 1009 | 23 | AAE19158 | Human kinase polyp |
| 385 | 7 | 0.8 | 630 | 22 | ABB67396 | Drosophila melanog | 458 | 7 | 0.8 | 1014 | 20 | AAV38743 | Neisseria meningit |
| 386 | 7 | 0.8 | 631 | 22 | ABB66645 | Drosophila melanog | 459 | 7 | 0.8 | 1015 | 20 | AAV38744 | Neisseria meningit |
| 387 | 7 | 0.8 | 631 | 23 | AAU12044 | Clostridium diffic | 460 | 7 | 0.8 | 1022 | 23 | AAU76875 | Human EphA extende |
| 388 | 7 | 0.8 | 632 | 23 | ABB49748 | Listeria monocytog | 461 | 7 | 0.8 | 1030 | 23 | ABP35638 | Fungal ZBC protein |
| 389 | 7 | 0.8 | 632 | 22 | ABB65156 | Drosophila melanog | 462 | 7 | 0.8 | 1032 | 23 | AAE16107 | Human DNAX Toll li |
| 390 | 7 | 0.8 | 634 | 22 | ABB07503 | Human GRP-binding | 463 | 7 | 0.8 | 1034 | 22 | AAU34738 | E. coli cellular p |
| 391 | 7 | 0.8 | 639 | 22 | ABG25372 | Novel human diagno | 464 | 7 | 0.8 | 1036 | 21 | AAV82776 | Human chordin rela |
| 392 | 7 | 0.8 | 641 | 23 | AAW49040 | Human testicular d | 465 | 7 | 0.8 | 1036 | 21 | AAV53034 | Human secreted pro |
| 393 | 7 | 0.8 | 641 | 24 | AAO16116 | Human cancer-testi | 466 | 7 | 0.8 | 1036 | 22 | AAU07141 | Human CRIM1 protei |
| 394 | 7 | 0.8 | 647 | 22 | ABB60758 | Drosophila melanog | 467 | 7 | 0.8 | 1036 | 22 | AAU12242 | Human PRO4330 poly |
| 395 | 7 | 0.8 | 654 | 23 | ABB54057 | Lactococcus lactis | 468 | 7 | 0.8 | 1036 | 23 | AAE18852 | Human pharmaceutical |
| 396 | 7 | 0.8 | 660 | 23 | ABB94293 | Chlamydia trachoma | 469 | 7 | 0.8 | 1036 | 24 | ABU66840 | Human PRO polypept |
| 397 | 7 | 0.8 | 661 | 22 | ABG30261 | Novel human diagno | 470 | 7 | 0.8 | 1036 | 24 | ABU66916 | Human secreted and |
| 398 | 7 | 0.8 | 670 | 23 | ABP26115 | Streptococcus poly | 471 | 7 | 0.8 | 1036 | 22 | ABU59721 | Novel secreted and |
| 399 | 7 | 0.8 | 671 | 21 | AAU07763 | Amino acid sequenc | 472 | 7 | 0.8 | 1040 | 22 | ABB23027 | Protein #5026 enco |
| 400 | 7 | 0.8 | 672 | 21 | AAU12123 | Corn MPPI. Zea ma | 473 | 7 | 0.8 | 1040 | 22 | AAW31141 | Peptide #5178 enco |
| 401 | 7 | 0.8 | 674 | 22 | AAU14364 | Human novel protei | 474 | 7 | 0.8 | 1040 | 23 | ABG40504 | Human peptide enco |
| 402 | 7 | 0.8 | 674 | 22 | AAU95048 | Human protein sequ | 475 | 7 | 0.8 | 1041 | 20 | AAV41168 | Human PRO286 prote |
| 403 | 7 | 0.8 | 674 | 23 | ABB76698 | Human eukaryotic i | 476 | 7 | 0.8 | 1041 | 20 | AAV05867 | Human Toll protein |
| 404 | 7 | 0.8 | 687 | 24 | ABB98844 | Human NEPHA-SV1.PR | 477 | 7 | 0.8 | 1041 | 21 | AAW44324 | Human PRO286 prote |
| 405 | 7 | 0.8 | 687 | 24 | ABB98851 | Human Eph receptor | 478 | 7 | 0.8 | 1041 | 23 | AAO21584 | Murine Toll-like r |
| 406 | 7 | 0.8 | 689 | 23 | ABB49714 | Listeria monocytog | 479 | 7 | 0.8 | 1041 | 23 | AAO21586 | Murine Toll-like r |
| 407 | 7 | 0.8 | 710 | 22 | AAU33871 | Staphylococcus aur | 480 | 7 | 0.8 | 1041 | 24 | ABU61154 | Human PRO286 polyp |
| 408 | 7 | 0.8 | 711 | 19 | AAW44842 | Staphylococcus aur | 481 | 7 | 0.8 | 1059 | 23 | AAO21587 | Murine Toll-like r |
| 409 | 7 | 0.8 | 711 | 22 | AAU36597 | Staphylococcus aur | 482 | 7 | 0.8 | 1070 | 21 | AAV93308 | A manganese supero |
| 410 | 7 | 0.8 | 732 | 22 | AAW61140 | Human NOV10 protei | 483 | 7 | 0.8 | 1070 | 22 | ABG12878 | Novel human diagno |
| 411 | 7 | 0.8 | 737 | 24 | ABB98847 | Human NEPHA-SV4.PR | 484 | 7 | 0.8 | 1075 | 19 | AAW64472 | Human secreted pro |
| 412 | 7 | 0.8 | 737 | 24 | ABB98854 | Human Eph receptor | 485 | 7 | 0.8 | 1075 | 22 | AAV90748 | Human DL162 2 pro |
| 413 | 7 | 0.8 | 754 | 21 | AAV35987 | Sorbitol dehydroge | 486 | 7 | 0.8 | 1083 | 20 | AAV25169 | Human RSC ligase p |
| 414 | 7 | 0.8 | 772 | 19 | AAW29666 | Homo sapiens DH130 | 487 | 7 | 0.8 | 1095 | 21 | AAV93311 | A manganese supero |
| 415 | 7 | 0.8 | 776 | 22 | ABG28762 | Novel human diagno | 488 | 7 | 0.8 | 1109 | 22 | AAW84930 | Shrimp white spot |
| 416 | 7 | 0.8 | 781 | 16 | AAW08986 | Human parvovirus V | 489 | 7 | 0.8 | 1109 | 20 | AYO23776 | Polypeptide identi |
| 417 | 7 | 0.8 | 781 | 20 | AAV23227 | Erythrovirus V9 VP | 490 | 7 | 0.8 | 1138 | 11 | AAO06461 | BtGS1245 protoxin |
| 418 | 7 | 0.8 | 781 | 24 | ABP57263 | Human parvovirus B | 491 | 7 | 0.8 | 1138 | 14 | AAO37213 | B.t. toxin HD511. |
| 419 | 7 | 0.8 | 781 | 24 | ABP57266 | Human parvovirus B | 492 | 7 | 0.8 | 1138 | 14 | AAO37214 | B.t. toxin HD867. |
| 420 | 7 | 0.8 | 784 | 22 | ABB63844 | Drosophila melanog | 493 | 7 | 0.8 | 1138 | 15 | AAW46225 | Bacillus thuringie |
| 421 | 7 | 0.8 | 812 | 22 | AAW78604 | Lawsonia intracell | 494 | 7 | 0.8 | 1138 | 15 | AAW46226 | Bacillus thuringie |
| 422 | 7 | 0.8 | 814 | 22 | ABB68374 | Drosophila melanog | 495 | 7 | 0.8 | 1160 | 22 | ABG28234 | C glutamicum prote |
| 423 | 7 | 0.8 | 817 | 14 | AAW34942 | Human epidermal tr | 496 | 7 | 0.8 | 1247 | 22 | AAV90997 | Novel human diagno |
| 424 | 7 | 0.8 | 817 | 23 | AAW49673 | Human TCK protein. | 497 | 7 | 0.8 | 1275 | 24 | ABU25467 | Aspergillus fumiga |
| 425 | 7 | 0.8 | 831 | 24 | ABP78978 | N. gonorrhoeae ami | 498 | 7 | 0.8 | 1287 | 16 | AAV79944 | Helicobacter pylor |
| 426 | 7 | 0.8 | 833 | 22 | ABB68116 | Drosophila melanog | 499 | 7 | 0.8 | 1288 | 18 | AAW55547 | H. pylori ORF 14ee |
| 427 | 7 | 0.8 | 836 | 22 | AAW62026 | Recombinant P. fur | 500 | 7 | 0.8 | 1288 | 18 | AAW55547 | H. pylori ORF 07ee |
| 428 | 7 | 0.8 | 839 | 22 | ABB63119 | Drosophila melanog | 501 | 7 | 0.8 | 1290 | 19 | AAW55865 | H. pylori GRP 374 |
| 429 | 7 | 0.8 | 843 | 22 | ABB69333 | Drosophila melanog | 502 | 7 | 0.8 | 1291 | 19 | AAW68392 | Clostridium botuli |
| 430 | 7 | 0.8 | 848 | 22 | AAW04082 | Botulinum toxin hea | 503 | 7 | 0.8 | 1296 | 14 | AAW41198 | CT. Helicobacter |
| 431 | 7 | 0.8 | 855 | 22 | AAW96787 | Putative P. abyssi | 504 | 7 | 0.8 | 1400 | 22 | AAW78576 | Human protein SEQ |
| 432 | 7 | 0.8 | 860 | 22 | ABG20680 | Novel human diagno | 505 | 7 | 0.8 | 1400 | 22 | AAW79560 | Human protein SEQ |
| 433 | 7 | 0.8 | 865 | 22 | ABG25368 | Novel human diagno | 506 | 7 | 0.8 | 1435 | 24 | ABU26067 | Aspergillus fumiga |
| 434 | 7 | 0.8 | 866 | 22 | ABB71210 | Drosophila melanog | 507 | 7 | 0.8 | 1476 | 22 | ABG58706 | Drosophila melanog |
| 435 | 7 | 0.8 | 868 | 22 | ABG30260 | Novel human diagno | 508 | 7 | 0.8 | 1497 | 22 | ABG20797 | Novel human diagno |
| 436 | 7 | 0.8 | 872 | 22 | ABB70960 | Drosophila melanog | 509 | 7 | 0.8 | 1600 | 22 | AAW84976 | Shrimp white spot |
| 437 | 7 | 0.8 | 873 | 23 | AAE25018 | Human drug metabol | 510 | 7 | 0.8 | 1605 | 22 | ABW70375 | Drosophila melanog |
| 438 | 7 | 0.8 | 875 | 24 | ABP96075 | Human protein kina | 511 | 7 | 0.8 | 1841 | 21 | AAW07562 | Protein encoded by |
| 439 | 7 | 0.8 | 877 | 24 | ABP96076 | Human protein kina | 512 | 7 | 0.8 | 1908 | 24 | ABU25817 | Aspergillus fumiga |
| 440 | 7 | 0.8 | 888 | 22 | ABB61813 | Drosophila melanog | 513 | 7 | 0.8 | 2022 | 24 | ABU26417 | Aspergillus fumiga |
| 441 | 7 | 0.8 | 892 | 22 | AAW62179 | Human p100 protein | 514 | 7 | 0.8 | 2099 | 24 | ABU25632 | Aspergillus fumiga |
| 442 | 7 | 0.8 | 893 | 22 | AAW67396 | Amino acid sequenc | 515 | 7 | 0.8 | 2162 | 21 | AAW07560 | Protein encoded by |
| 443 | 7 | 0.8 | 893 | 22 | AAW98890 | Novel human (NHP) | 516 | 7 | 0.8 | 2405 | 24 | ABU26232 | Aspergillus fumiga |
| 444 | 7 | 0.8 | 893 | 23 | AAU76324 | Human protein kina | 517 | 7 | 0.8 | 2517 | 22 | ABG65918 | Drosophila melanog |
| 445 | 7 | 0.8 | 898 | 21 | AAV83939 | Benzene ring hydro | 518 | 7 | 0.8 | 2519 | 22 | ABG16636 | Novel human diagno |
| 446 | 7 | 0.8 | 899 | 22 | ABG62277 | Drosophila melanog | 519 | 7 | 0.8 | 3010 | 16 | AAW95022 | Hepatitis GB virus |
| 447 | 7 | 0.8 | 929 | 22 | ABB70857 | Drosophila melanog | 520 | 7 | 0.8 | 3135 | 15 | AAW57474 | P. falciparum tran |

| | | | | | | | | | | | | |
|-----|---|-----|----------|----|----------|---------------------|-----|-----|----|----|-----------|--------------------|
| 521 | 7 | 0.8 | 3135 | 21 | AAB18223 | Plasmodium falcipar | 594 | 0.6 | 21 | 22 | AAG65544 | Antimicrobial pept |
| 522 | 7 | 0.8 | 3234 | 23 | ABP70132 | Human NOV39a Hom | 595 | 0.6 | 21 | 23 | ABG61449 | Tick Ixolaris seco |
| 523 | 7 | 0.8 | ABP70847 | 24 | AAB70847 | Amino acid sequenc | 596 | 0.6 | 22 | 23 | ABG61448 | Tick Ixolaris seco |
| 524 | 7 | 0.8 | 3433 | 24 | ABP98821 | West Nile virus st | 597 | 0.6 | 22 | 23 | AAU69693 | Cell death protect |
| 525 | 7 | 0.8 | 4999 | 23 | AAO22158 | Ramoplanin biosynt | 598 | 0.6 | 23 | 21 | AAI69795 | Human interleukin- |
| 526 | 7 | 0.8 | 19938 | 24 | ABB98398 | Streptomyces virid | 599 | 0.6 | 23 | 23 | ABG61447 | Tick Ixolaris seco |
| 527 | 6 | 0.6 | 8 | 22 | ABP78491 | SIV and HIV gp 41 | 600 | 0.6 | 23 | 23 | ABBS3201 | Human ORF107 prote |
| 528 | 6 | 0.6 | 9 | 22 | AAU24331 | Human MHC class I | 601 | 0.6 | 24 | 13 | AAR27088 | OmpF signal peptid |
| 529 | 6 | 0.6 | 9 | 22 | AAU24443 | Human MHC molecule | 602 | 0.6 | 24 | 15 | AAR50167 | OmpF leader sequen |
| 530 | 6 | 0.6 | 9 | 22 | AAE26784 | Fibrin binding pep | 603 | 0.6 | 24 | 18 | AAW11712 | E.coli leader sequ |
| 531 | 6 | 0.6 | 10 | 20 | AAI46794 | Immunogenic peptid | 604 | 0.6 | 24 | 22 | AAE03613 | Human AZAD protein |
| 532 | 6 | 0.6 | 10 | 20 | AAI07312 | Collagen assembly | 605 | 0.6 | 24 | 22 | AAAB2365 | Protein sequence s |
| 533 | 6 | 0.6 | 10 | 22 | AAU23878 | Human MHC class I | 606 | 0.6 | 24 | 23 | ABB76468 | Synthetic leader p |
| 534 | 6 | 0.6 | 10 | 22 | AAU23882 | Human MHC class I | 607 | 0.6 | 24 | 23 | ABB76469 | Synthetic leader p |
| 535 | 6 | 0.6 | 10 | 22 | AAU24213 | Human MHC molecule | 608 | 0.6 | 24 | 23 | ABB76472 | Synthetic leader p |
| 536 | 6 | 0.6 | 10 | 22 | AAI95769 | Human complementar | 609 | 0.6 | 24 | 23 | ABG61446 | Tick Ixolaris seco |
| 537 | 6 | 0.6 | 10 | 22 | AAI84417 | Arabidopsis thalia | 610 | 0.6 | 24 | 23 | AAU69692 | Cell death protect |
| 538 | 6 | 0.6 | 10 | 22 | AAI86218 | Saccharomyces cere | 611 | 0.6 | 25 | 21 | AAI80001 | Optimised IGR-CH3 |
| 539 | 6 | 0.6 | 10 | 22 | AAI86454 | Saccharomyces cere | 612 | 0.6 | 25 | 23 | ABG61445 | Tick Ixolaris seco |
| 540 | 6 | 0.6 | 10 | 23 | ABG98131 | Anti-neovascular p | 613 | 0.6 | 26 | 12 | AAR12451 | HIBP1-9(307-331) H |
| 541 | 6 | 0.6 | 10 | 24 | ABU03336 | Human expressed pr | 614 | 0.6 | 26 | 14 | AAR40061 | Hib OMP P1 peptide |
| 542 | 6 | 0.6 | 10 | 24 | ABP74549 | Human PSMA epitope | 615 | 0.6 | 26 | 15 | AAR54955 | Antoi grass pollen |
| 543 | 6 | 0.6 | 12 | 15 | AAR60404 | Antiproliferative | 616 | 0.6 | 26 | 21 | AAI84094 | Synthetic amino ac |
| 544 | 6 | 0.6 | 12 | 22 | AAU28915 | MART-1 antigen pep | 617 | 0.6 | 26 | 21 | AAI86459 | Human gene 47-enco |
| 545 | 6 | 0.6 | 13 | 15 | AAI43970 | N-heparan sulphate | 618 | 0.6 | 26 | 22 | ABG57550 | Human liver peptid |
| 546 | 6 | 0.6 | 13 | 23 | ABP79454 | Alanine-stabilised | 619 | 0.6 | 26 | 22 | ABB42120 | Peptide #9626 enco |
| 547 | 6 | 0.6 | 13 | 23 | ABP79457 | BDC2.5 peptide. U | 620 | 0.6 | 26 | 22 | AAI63004 | Human bone marrow |
| 548 | 6 | 0.6 | 14 | 22 | ABBS6594 | Human SNP related | 621 | 0.6 | 26 | 22 | AAI75815 | Human immune/haema |
| 549 | 6 | 0.6 | 15 | 20 | AAI07417 | HSV-1 TK amino aci | 622 | 0.6 | 26 | 22 | AAI88302 | Peptide #9960 enco |
| 550 | 6 | 0.6 | 15 | 20 | AAI07309 | Collagen assembly | 623 | 0.6 | 26 | 22 | AAI35923 | Tick Ixolaris seco |
| 551 | 6 | 0.6 | 15 | 20 | AAI97149 | HSV-1 TK amino aci | 624 | 0.6 | 26 | 23 | ABG61444 | Tick Ixolaris seco |
| 552 | 6 | 0.6 | 15 | 21 | AAI15883 | Human chemokine de | 625 | 0.6 | 27 | 23 | ABG61443 | Tick Ixolaris seco |
| 553 | 6 | 0.6 | 15 | 21 | AAI14218 | Mutant HIV cleavag | 626 | 0.6 | 28 | 18 | AAI12665 | Orange pectin meth |
| 554 | 6 | 0.6 | 15 | 23 | ABP59092 | Cell division cycl | 627 | 0.6 | 28 | 21 | AAI84088 | Amino acid sequenc |
| 555 | 6 | 0.6 | 15 | 23 | AAI79657 | Development regula | 628 | 0.6 | 28 | 22 | ABG49076 | Human liver peptid |
| 556 | 6 | 0.6 | 15 | 23 | ABP79475 | Fibrin binding pep | 629 | 0.6 | 28 | 22 | ABG58481 | Human liver peptid |
| 557 | 6 | 0.6 | 15 | 23 | ABP79475 | Human Orc4p 47.3 | 630 | 0.6 | 28 | 22 | ABB29077 | Peptide #1728 enco |
| 558 | 6 | 0.6 | 15 | 23 | ABG65786 | Plant potential PE | 631 | 0.6 | 28 | 22 | ABB34237 | Peptide #1743 enco |
| 559 | 6 | 0.6 | 15 | 23 | AAU77490 | N-terminal peptide | 632 | 0.6 | 28 | 22 | ABB43081 | Peptide #10587 enc |
| 560 | 6 | 0.6 | 15 | 23 | ABG99208 | Thymidine kinase (| 633 | 0.6 | 28 | 22 | ABB19673 | Protein #1672 enco |
| 561 | 6 | 0.6 | 16 | 18 | AAI12672 | Orange pectin meth | 634 | 0.6 | 28 | 22 | ABB26235 | Protein #8234 enco |
| 562 | 6 | 0.6 | 16 | 22 | AAI64741 | Human secreted pro | 635 | 0.6 | 28 | 22 | AAI55032 | Human brain expres |
| 563 | 6 | 0.6 | 16 | 22 | AAI11442 | Interleukin 12 der | 636 | 0.6 | 28 | 22 | AAI63390 | Human brain expres |
| 564 | 6 | 0.6 | 17 | 21 | AAI28076 | Human secreted pro | 637 | 0.6 | 28 | 22 | AAI67419 | Human bone marrow |
| 565 | 6 | 0.6 | 17 | 23 | AAI23446 | Mouse and rat STIM | 638 | 0.6 | 28 | 22 | AAI76800 | Human bone marrow |
| 566 | 6 | 0.6 | 17 | 23 | AAI23447 | Mouse STIM2 peptid | 639 | 0.6 | 28 | 22 | AAI15247 | Peptide #1681 enco |
| 567 | 6 | 0.6 | 18 | 11 | AAI06540 | Human B-cell simul | 640 | 0.6 | 28 | 22 | AAI27708 | Peptide #1745 enco |
| 568 | 6 | 0.6 | 18 | 16 | AAI70727 | Signal peptide use | 641 | 0.6 | 28 | 22 | AAI36906 | Peptide #10943 enc |
| 569 | 6 | 0.6 | 18 | 16 | AAI76687 | Signal peptide use | 642 | 0.6 | 28 | 22 | AAI02991 | Peptide #1673 enco |
| 570 | 6 | 0.6 | 18 | 23 | AAE28328 | Human Dinp protein | 643 | 0.6 | 28 | 23 | ABG37042 | Human peptid enco |
| 571 | 6 | 0.6 | 18 | 23 | ABG61452 | Tick Ixolaris seco | 644 | 0.6 | 28 | 23 | ABG45958 | Human peptid enco |
| 572 | 6 | 0.6 | 18 | 24 | ABJ38351 | Preferred TALL-1 m | 645 | 0.6 | 28 | 23 | ABG61442 | Tick Ixolaris seco |
| 573 | 6 | 0.6 | 18 | 24 | ABP54217 | Human lung specifi | 646 | 0.6 | 29 | 23 | ABG61441 | Tick Ixolaris seco |
| 574 | 6 | 0.6 | 19 | 16 | AAI86351 | HIV-1 protease inh | 647 | 0.6 | 30 | 4 | AAI30193 | Sequence encoded b |
| 575 | 6 | 0.6 | 19 | 16 | AAI86352 | HIV-1 protease inh | 648 | 0.6 | 30 | 16 | AAI75651 | Human placenta der |
| 576 | 6 | 0.6 | 19 | 19 | AAI23945 | Human herpesvirus | 649 | 0.6 | 30 | 17 | AAI92064 | Recombinant alpha2 |
| 577 | 6 | 0.6 | 19 | 23 | ABG61451 | Tick Ixolaris seco | 650 | 0.6 | 30 | 21 | AAI39084 | Human secreted pro |
| 578 | 6 | 0.6 | 20 | 12 | AAI10909 | Immunogenic epitop | 651 | 0.6 | 30 | 21 | AAI39084 | Human secreted pro |
| 579 | 6 | 0.6 | 20 | 12 | AAI14484 | LKT neutralising e | 652 | 0.6 | 30 | 22 | ABG15956 | Novel human thalia |
| 580 | 6 | 0.6 | 20 | 13 | AAI24125 | Leukotoxin epitope | 653 | 0.6 | 30 | 22 | AAU00948 | Human B7-H3 polype |
| 581 | 6 | 0.6 | 20 | 15 | AAI45183 | LKT synthetic epit | 654 | 0.6 | 30 | 22 | AAI887416 | Human secreted pro |
| 582 | 6 | 0.6 | 20 | 18 | AAI13868 | Synthetic leukotox | 655 | 0.6 | 30 | 22 | ABG61440 | Tick Ixolaris seco |
| 583 | 6 | 0.6 | 20 | 19 | AAI67508 | IGF-1/IGFBP inhibi | 656 | 0.6 | 31 | 22 | ABG50089 | Human liver peptid |
| 584 | 6 | 0.6 | 20 | 21 | AAI21076 | Pasteurella haemol | 657 | 0.6 | 31 | 22 | ABG30040 | Peptide #2691 enco |
| 585 | 6 | 0.6 | 20 | 21 | AAI84052 | Amino acid sequenc | 658 | 0.6 | 31 | 22 | ABG35211 | Peptide #2717 enco |
| 586 | 6 | 0.6 | 20 | 22 | AAI84052 | Synthetic peptide | 659 | 0.6 | 31 | 22 | ABG20652 | Protein #2651 enco |
| 587 | 6 | 0.6 | 20 | 23 | AAI28335 | Human Dinp protein | 660 | 0.6 | 31 | 22 | AAI56043 | Human brain expres |
| 588 | 6 | 0.6 | 20 | 23 | ABG31560 | Interleukin-12 pep | 661 | 0.6 | 31 | 22 | AAI68415 | Human bone marrow |
| 589 | 6 | 0.6 | 20 | 23 | ABG61450 | Tick Ixolaris seco | 662 | 0.6 | 31 | 22 | AAI16230 | Peptide #2664 enco |
| 590 | 6 | 0.6 | 20 | 23 | ABP57663 | IGFBP-3 binding pe | 663 | 0.6 | 31 | 22 | AAI28720 | Peptide #2757 enco |
| 591 | 6 | 0.6 | 21 | 16 | AAI76686 | Signal peptide use | 664 | 0.6 | 31 | 22 | AAI03960 | Peptide #2642 enco |
| 592 | 6 | 0.6 | 21 | 22 | AAI23787 | Neisseria gonorrhoe | 665 | 0.6 | 31 | 23 | ABG37988 | Human peptid enco |
| 593 | 6 | 0.6 | 21 | 22 | AAI65528 | Antimicrobial pept | 666 | 0.6 | 31 | 23 | ABG61439 | Tick Ixolaris seco |

| | | | | | | | | | | | | | |
|-----|---|-----|----|----|----------|--------------------|-----|---|-----|----|----|----------|--------------------|
| 667 | 6 | 0.6 | 32 | 23 | ABG61438 | Tick Ixolaris seco | 740 | 6 | 0.6 | 44 | 23 | ABG61426 | Tick Ixolaris seco |
| 668 | 6 | 0.6 | 32 | 23 | ABP28129 | Streptococcus poly | 741 | 6 | 0.6 | 44 | 24 | ABP79654 | N. gonorrhoeae ami |
| 669 | 6 | 0.6 | 32 | 21 | AAV50193 | Streptococcus pyog | 742 | 6 | 0.6 | 45 | 21 | AAV80018 | IgE immunogenic pe |
| 670 | 6 | 0.6 | 33 | 22 | ABG49558 | Human liver peptid | 743 | 6 | 0.6 | 45 | 21 | ABG54757 | Human liver peptid |
| 671 | 6 | 0.6 | 33 | 22 | ABG29950 | Peptide #2201 enco | 744 | 6 | 0.6 | 45 | 22 | ABG33689 | Peptide #7195 enco |
| 672 | 6 | 0.6 | 33 | 22 | ABG34734 | Peptide #2240 enco | 745 | 6 | 0.6 | 45 | 22 | ABG24351 | Protein #6350 enco |
| 673 | 6 | 0.6 | 33 | 22 | ABG20146 | Protein #2145 enco | 746 | 6 | 0.6 | 45 | 22 | ABG60407 | Human brain expres |
| 674 | 6 | 0.6 | 33 | 22 | AAV55532 | Human brain expres | 747 | 6 | 0.6 | 45 | 22 | AAV70731 | Human bone marrow |
| 675 | 6 | 0.6 | 33 | 22 | AAV67916 | Human bone marrow | 748 | 6 | 0.6 | 45 | 22 | AAV73043 | Human bone marrow |
| 676 | 6 | 0.6 | 33 | 22 | AAV15733 | Peptide #2167 enco | 749 | 6 | 0.6 | 45 | 22 | AAV19816 | Peptide #6250 enco |
| 677 | 6 | 0.6 | 33 | 22 | AAV28238 | Peptide #2275 enco | 750 | 6 | 0.6 | 45 | 22 | AAV33265 | Peptide #7302 enco |
| 678 | 6 | 0.6 | 33 | 22 | AAV03466 | Peptide #2148 enco | 751 | 6 | 0.6 | 45 | 22 | ABG66328 | D radiodurans glut |
| 679 | 6 | 0.6 | 33 | 22 | AAV48668 | Human CD80 constan | 752 | 6 | 0.6 | 45 | 23 | ABG40467 | Human peptid enco |
| 680 | 6 | 0.6 | 33 | 22 | ABG37452 | Human peptid enco | 753 | 6 | 0.6 | 45 | 23 | ABG42887 | Human peptid enco |
| 681 | 6 | 0.6 | 33 | 23 | ABG61437 | Tick Ixolaris seco | 754 | 6 | 0.6 | 45 | 23 | ABG61378 | Tick Ixolaris seco |
| 682 | 6 | 0.6 | 34 | 14 | AAV36607 | IL-6 receptor bind | 755 | 6 | 0.6 | 45 | 23 | ABG61425 | Tick Ixolaris seco |
| 683 | 6 | 0.6 | 34 | 33 | ABG61436 | Tick Ixolaris seco | 756 | 6 | 0.6 | 46 | 21 | AAV80017 | Caenorhabditis ele |
| 684 | 6 | 0.6 | 35 | 21 | AAV65241 | Human 5' EST relat | 757 | 6 | 0.6 | 46 | 21 | AAV80017 | IgE immunogenic pe |
| 685 | 6 | 0.6 | 35 | 23 | ABG70821 | Mouse myocardin as | 758 | 6 | 0.6 | 46 | 22 | ABG49456 | Human liver peptid |
| 686 | 6 | 0.6 | 35 | 23 | ABG61435 | Tick Ixolaris seco | 759 | 6 | 0.6 | 46 | 22 | ABG15316 | Novel human diagno |
| 687 | 6 | 0.6 | 35 | 24 | ABG74189 | Mouse myocardin as | 760 | 6 | 0.6 | 46 | 22 | ABG29456 | Peptide #2107 enco |
| 688 | 6 | 0.6 | 36 | 21 | AAV18727 | Amino acid sequenc | 761 | 6 | 0.6 | 46 | 22 | ABG34633 | Peptide #2139 enco |
| 689 | 6 | 0.6 | 36 | 22 | ABV40753 | Peptide #8259 enco | 762 | 6 | 0.6 | 46 | 22 | ABG20043 | Protein #2042 enco |
| 690 | 6 | 0.6 | 36 | 22 | AAV61613 | Human brain expres | 763 | 6 | 0.6 | 46 | 22 | AAV55423 | Human brain expres |
| 691 | 6 | 0.6 | 36 | 22 | AAV74405 | Human bone marrow | 764 | 6 | 0.6 | 46 | 22 | AAV67818 | Human bone marrow |
| 692 | 6 | 0.6 | 36 | 23 | ABJ11196 | Yeast selected int | 765 | 6 | 0.6 | 46 | 22 | AAV15628 | Peptide #2062 enco |
| 693 | 6 | 0.6 | 36 | 23 | ABG77294 | Selected Interacti | 766 | 6 | 0.6 | 46 | 22 | AAV28126 | Peptide #2163 enco |
| 694 | 6 | 0.6 | 36 | 23 | ABG61434 | Tick Ixolaris seco | 767 | 6 | 0.6 | 46 | 22 | AAV03371 | Peptide #2053 enco |
| 695 | 6 | 0.6 | 36 | 24 | ABP79856 | N. gonorrhoeae ami | 768 | 6 | 0.6 | 46 | 23 | ABG37361 | Human peptid enco |
| 696 | 6 | 0.6 | 37 | 22 | ABG53245 | Human liver peptid | 769 | 6 | 0.6 | 46 | 23 | ABG61377 | Tick Ixolaris seco |
| 697 | 6 | 0.6 | 37 | 22 | ABG38423 | Peptide #5929 enco | 770 | 6 | 0.6 | 46 | 23 | ABG61424 | Tick Ixolaris seco |
| 698 | 6 | 0.6 | 37 | 22 | AAV59033 | Human brain expres | 771 | 6 | 0.6 | 47 | 21 | AAV46731 | Arabidopsis thalia |
| 699 | 6 | 0.6 | 37 | 22 | AAV71563 | Human bone marrow | 772 | 6 | 0.6 | 47 | 23 | ABG61376 | Tick Ixolaris seco |
| 700 | 6 | 0.6 | 37 | 22 | AAV31861 | Peptide #5898 enco | 773 | 6 | 0.6 | 47 | 23 | ABG61423 | Tick Ixolaris seco |
| 701 | 6 | 0.6 | 37 | 23 | ABG41374 | Human peptid enco | 774 | 6 | 0.6 | 48 | 14 | AAV34858 | Peptide showing CA |
| 702 | 6 | 0.6 | 37 | 23 | ABG61433 | Tick Ixolaris seco | 775 | 6 | 0.6 | 48 | 20 | AAV88643 | Secreted protein e |
| 703 | 6 | 0.6 | 37 | 24 | ABV47947 | Human secreted pro | 776 | 6 | 0.6 | 48 | 22 | ABV50410 | Human secreted pro |
| 704 | 6 | 0.6 | 37 | 24 | ABR00179 | Human gene 169 enc | 777 | 6 | 0.6 | 48 | 22 | AAV99015 | Wheat 33K signal p |
| 705 | 6 | 0.6 | 38 | 21 | AAV01388 | Secreted protein e | 778 | 6 | 0.6 | 48 | 23 | ABG61375 | Tick Ixolaris seco |
| 706 | 6 | 0.6 | 38 | 21 | AAV56549 | Human prostate can | 779 | 6 | 0.6 | 48 | 23 | ABG61422 | Tick Ixolaris seco |
| 707 | 6 | 0.6 | 38 | 23 | AAE28324 | Human DinP protein | 780 | 6 | 0.6 | 49 | 23 | AAV87386 | Human gene 45 enco |
| 708 | 6 | 0.6 | 38 | 23 | ABG61432 | Tick Ixolaris seco | 781 | 6 | 0.6 | 49 | 23 | ABG65353 | Human albumin fusi |
| 709 | 6 | 0.6 | 38 | 23 | ABP30565 | Streptococcus poly | 782 | 6 | 0.6 | 49 | 23 | ABG61374 | Tick Ixolaris seco |
| 710 | 6 | 0.6 | 38 | 24 | ABP96039 | Mus musculus DGAT | 783 | 6 | 0.6 | 49 | 23 | ABG61421 | Tick Ixolaris seco |
| 711 | 6 | 0.6 | 39 | 21 | AAV65068 | Human 5' EST relat | 784 | 6 | 0.6 | 50 | 20 | AAV33423 | L. brevis ADH pept |
| 712 | 6 | 0.6 | 39 | 22 | AAV20686 | Human novel foetal | 785 | 6 | 0.6 | 50 | 22 | AAO21254 | Interleukin-2 homo |
| 713 | 6 | 0.6 | 39 | 23 | ABG61384 | Tick Ixolaris seco | 786 | 6 | 0.6 | 50 | 22 | AAU62599 | Propionibacterium |
| 714 | 6 | 0.6 | 39 | 23 | ABG61431 | Tick Ixolaris seco | 787 | 6 | 0.6 | 50 | 22 | AAO11706 | Human polypeptide |
| 715 | 6 | 0.6 | 39 | 24 | ABP96500 | Peptide-linker-bet | 788 | 6 | 0.6 | 50 | 22 | AAV77522 | Human colon cancer |
| 716 | 6 | 0.6 | 40 | 22 | AAV99008 | Spinach 23K transf | 789 | 6 | 0.6 | 50 | 23 | ABG61373 | Tick Ixolaris seco |
| 717 | 6 | 0.6 | 40 | 23 | ABG61383 | Tick Ixolaris seco | 790 | 6 | 0.6 | 50 | 23 | ABG61420 | Tick Ixolaris seco |
| 718 | 6 | 0.6 | 40 | 23 | ABG61430 | Tick Ixolaris seco | 791 | 6 | 0.6 | 51 | 21 | AAV34384 | Arabidopsis thalia |
| 719 | 6 | 0.6 | 41 | 19 | AAV42440 | S. pombe MEm1 mat1 | 792 | 6 | 0.6 | 51 | 23 | ABG61372 | Tick Ixolaris seco |
| 720 | 6 | 0.6 | 41 | 20 | AAV14018 | Human beta-catenin | 793 | 6 | 0.6 | 51 | 23 | ABP11189 | Human ORFX protein |
| 721 | 6 | 0.6 | 41 | 23 | ABG61382 | Tick Ixolaris seco | 794 | 6 | 0.6 | 52 | 22 | ABG49785 | Human liver peptid |
| 722 | 6 | 0.6 | 41 | 23 | ABG61429 | Tick Ixolaris seco | 795 | 6 | 0.6 | 52 | 22 | AAV58374 | Propionibacterium |
| 723 | 6 | 0.6 | 42 | 23 | ABG61381 | Tick Ixolaris seco | 796 | 6 | 0.6 | 52 | 22 | AAV55757 | Human brain expres |
| 724 | 6 | 0.6 | 42 | 23 | ABG61428 | Tick Ixolaris seco | 797 | 6 | 0.6 | 52 | 22 | AAV68139 | Human bone marrow |
| 725 | 6 | 0.6 | 43 | 20 | AAV14018 | YhdD repeat sequen | 798 | 6 | 0.6 | 53 | 18 | AAV55593 | H. pylori ORF 09cp |
| 726 | 6 | 0.6 | 43 | 22 | ABG53412 | Human liver peptid | 799 | 6 | 0.6 | 53 | 20 | AAV12475 | Human 5' EST secre |
| 727 | 6 | 0.6 | 43 | 22 | ABG00780 | Novel human diagno | 800 | 6 | 0.6 | 53 | 21 | AAV28752 | Sequence homologou |
| 728 | 6 | 0.6 | 43 | 22 | ABG11360 | Novel human diagno | 801 | 6 | 0.6 | 53 | 22 | AAO21245 | Novel interleukin- |
| 729 | 6 | 0.6 | 43 | 22 | ABV38568 | Peptide #6074 enco | 802 | 6 | 0.6 | 53 | 22 | AAO21246 | Interleukin-2 homo |
| 730 | 6 | 0.6 | 43 | 22 | AAV59192 | Human brain expres | 803 | 6 | 0.6 | 53 | 22 | AAO11297 | Human polypeptide |
| 731 | 6 | 0.6 | 43 | 22 | AAV71727 | Human bone marrow | 804 | 6 | 0.6 | 53 | 22 | AAV74943 | Human colon cancer |
| 732 | 6 | 0.6 | 43 | 22 | AAV32019 | Peptide #6056 enco | 805 | 6 | 0.6 | 53 | 22 | ABG61190 | Tick Ixolaris amin |
| 733 | 6 | 0.6 | 43 | 23 | AAE28333 | Human DinP protein | 806 | 6 | 0.6 | 53 | 24 | ABP77646 | N. gonorrhoeae ami |
| 734 | 6 | 0.6 | 43 | 23 | ABG41541 | Human peptid enco | 807 | 6 | 0.6 | 54 | 22 | AAO21247 | Interleukin-2 homo |
| 735 | 6 | 0.6 | 43 | 23 | ABG61380 | Tick Ixolaris seco | 808 | 6 | 0.6 | 54 | 22 | AAO21248 | Interleukin-2 homo |
| 736 | 6 | 0.6 | 43 | 23 | ABG61427 | Tick Ixolaris seco | 809 | 6 | 0.6 | 54 | 22 | AAO21249 | Interleukin-2 homo |
| 737 | 6 | 0.6 | 44 | 21 | AAV24500 | Human secreted pro | 810 | 6 | 0.6 | 54 | 22 | AAO21251 | Interleukin-2 homo |
| 738 | 6 | 0.6 | 44 | 22 | AAV90069 | Human immune/haema | 811 | 6 | 0.6 | 54 | 22 | AAO21252 | Interleukin-2 homo |
| 739 | 6 | 0.6 | 44 | 23 | ABG61379 | Tick Ixolaris seco | 812 | 6 | 0.6 | 54 | 22 | AAU44002 | Propionibacterium |

| | | | | | | | | | | | | | |
|-----|---|-----|----|----|-----------|----------------------|-----|---|-----|----|----|-----------|---------------------|
| 813 | 6 | 0.6 | 54 | 22 | AAU54145 | Propionibacterium | 886 | 6 | 0.6 | 65 | 23 | AAO22521 | Small acid-soluble |
| 814 | 6 | 0.6 | 54 | 22 | AAU86435 | Human immune/haema | 887 | 6 | 0.6 | 65 | 23 | AAO22521 | Tick ixolaris amin |
| 815 | 6 | 0.6 | 54 | 23 | ABP62961 | Human polypeptide | 888 | 6 | 0.6 | 66 | 22 | AAU56053 | Propionibacterium |
| 816 | 6 | 0.6 | 54 | 23 | ABG61191 | Tick ixolaris amin | 889 | 6 | 0.6 | 66 | 22 | ABG16373 | Novel human diagno |
| 817 | 6 | 0.6 | 54 | 23 | ABP34785 | Human ORF3758 prot | 890 | 6 | 0.6 | 66 | 23 | ABE28316 | Human Dinp protein |
| 818 | 6 | 0.6 | 54 | 23 | ABP01675 | Human ORFX protein | 891 | 6 | 0.6 | 66 | 23 | ABG61203 | Tick ixolaris amin |
| 819 | 6 | 0.6 | 54 | 23 | ABP04933 | Human ORFX protein | 892 | 6 | 0.6 | 66 | 23 | ABP33629 | Human dehydrogenas |
| 820 | 6 | 0.6 | 54 | 23 | ABP06534 | Human ORFX protein | 893 | 6 | 0.6 | 67 | 23 | ABG61204 | Tick ixolaris amin |
| 821 | 6 | 0.6 | 55 | 22 | AAU45099 | Propionibacterium | 894 | 6 | 0.6 | 67 | 23 | ABP31120 | Human ORF93 protei |
| 822 | 6 | 0.6 | 55 | 22 | AAU18585 | Human lung antigen | 895 | 6 | 0.6 | 67 | 23 | ABP09599 | Human ORFX protein |
| 823 | 6 | 0.6 | 55 | 22 | AAU17855 | Novel human respir | 896 | 6 | 0.6 | 68 | 22 | ABE17375 | Human nervous syst |
| 824 | 6 | 0.6 | 55 | 23 | ABG61192 | Tick ixolaris amin | 897 | 6 | 0.6 | 68 | 23 | ABP21205 | Tick ixolaris amin |
| 825 | 6 | 0.6 | 56 | 22 | ABG78035 | Piscirickettsia sa | 898 | 6 | 0.6 | 68 | 23 | ABP25929 | Streptococcus poly |
| 826 | 6 | 0.6 | 56 | 23 | ABG61193 | Tick ixolaris amin | 899 | 6 | 0.6 | 69 | 22 | ABG56480 | Human liver peptid |
| 827 | 6 | 0.6 | 57 | 22 | ABE12132 | Human bone marrow | 900 | 6 | 0.6 | 69 | 22 | ABE41037 | Peptide #8543 enco |
| 828 | 6 | 0.6 | 57 | 23 | ABG61194 | Tick ixolaris amin | 901 | 6 | 0.6 | 69 | 22 | ABM61894 | Human brain expres |
| 829 | 6 | 0.6 | 58 | 22 | AAU41360 | Propionibacterium | 902 | 6 | 0.6 | 69 | 22 | AAU74896 | Human bone marrow |
| 830 | 6 | 0.6 | 58 | 22 | AAO08269 | Human polypeptide | 903 | 6 | 0.6 | 69 | 22 | AAU34812 | Peptide #8849 enco |
| 831 | 6 | 0.6 | 58 | 23 | ABG61195 | Tick ixolaris amin | 904 | 6 | 0.6 | 69 | 22 | AAE91815 | C Glutamicum prote |
| 832 | 6 | 0.6 | 58 | 23 | ABP03070 | Human ORFX protein | 905 | 6 | 0.6 | 69 | 23 | AAE28306 | Human Dinp protein |
| 833 | 6 | 0.6 | 58 | 23 | ABP03723 | Human ORFX protein | 906 | 6 | 0.6 | 69 | 23 | ABG44507 | Human peptide enco |
| 834 | 6 | 0.6 | 59 | 21 | AAU33675 | Microfibril associ | 907 | 6 | 0.6 | 69 | 23 | ABG61206 | Tick ixolaris amin |
| 835 | 6 | 0.6 | 59 | 21 | AAU76330 | Fragment of human | 908 | 6 | 0.6 | 70 | 21 | AAO2211 | Human secreted pro |
| 836 | 6 | 0.6 | 59 | 22 | AAU60996 | Propionibacterium | 909 | 6 | 0.6 | 70 | 21 | AAO2211 | Human secreted pro |
| 837 | 6 | 0.6 | 59 | 22 | ABG26902 | Novel human diagno | 910 | 6 | 0.6 | 70 | 22 | AAU40286 | Propionibacterium |
| 838 | 6 | 0.6 | 59 | 22 | AAO11340 | Human polypeptide | 911 | 6 | 0.6 | 70 | 22 | AAU54423 | Propionibacterium |
| 839 | 6 | 0.6 | 59 | 23 | ABP65003 | Human protein SEQ | 912 | 6 | 0.6 | 70 | 22 | ABG23797 | Novel human diagno |
| 840 | 6 | 0.6 | 59 | 23 | ABG61196 | Tick ixolaris amin | 913 | 6 | 0.6 | 70 | 22 | AAU94363 | Human reproductive |
| 841 | 6 | 0.6 | 60 | 19 | AAW62252 | Subtilase HSFURI f | 914 | 6 | 0.6 | 70 | 22 | AAU24031 | Murine EST encoded |
| 842 | 6 | 0.6 | 60 | 21 | ABO3674 | Microfibrillar aas | 915 | 6 | 0.6 | 70 | 22 | AAU42254 | Human breast or ov |
| 843 | 6 | 0.6 | 60 | 21 | AAO1532 | Human secreted pro | 916 | 6 | 0.6 | 70 | 23 | ABG61207 | Tick ixolaris amin |
| 844 | 6 | 0.6 | 60 | 22 | AAU51783 | Propionibacterium | 917 | 6 | 0.6 | 71 | 20 | AAU01550 | Antigenic epitope |
| 845 | 6 | 0.6 | 60 | 22 | AAU64475 | Propionibacterium | 918 | 6 | 0.6 | 71 | 20 | AAU12033 | Human 5' EST secre |
| 846 | 6 | 0.6 | 60 | 23 | ABG61197 | Tick ixolaris amin | 919 | 6 | 0.6 | 71 | 23 | ABG61208 | Tick ixolaris amin |
| 847 | 6 | 0.6 | 60 | 23 | ABE57406 | Human secreted pro | 920 | 6 | 0.6 | 71 | 24 | ABG76380 | Translation domain |
| 848 | 6 | 0.6 | 61 | 22 | AAU48693 | Propionibacterium | 921 | 6 | 0.6 | 72 | 21 | ABE333181 | Eucalyptus grandis |
| 849 | 6 | 0.6 | 61 | 22 | AAU22386 | Human cardiovascular | 922 | 6 | 0.6 | 72 | 21 | ABE38336 | Human secreted pro |
| 850 | 6 | 0.6 | 61 | 22 | AAU90453 | Human immune/haema | 923 | 6 | 0.6 | 72 | 22 | AAU54405 | Propionibacterium |
| 851 | 6 | 0.6 | 61 | 23 | ABU51140 | Helicobacter pylor | 924 | 6 | 0.6 | 72 | 22 | AAU94315 | Human reproductive |
| 852 | 6 | 0.6 | 61 | 23 | ABG61198 | Tick ixolaris amin | 925 | 6 | 0.6 | 72 | 22 | ABE80421 | Gene #6 associated |
| 853 | 6 | 0.6 | 61 | 23 | ABU39427 | Staphylococcus epi | 926 | 6 | 0.6 | 72 | 23 | ABG61209 | Tick ixolaris amin |
| 854 | 6 | 0.6 | 61 | 24 | ABU00458 | S. pneumoniae type | 927 | 6 | 0.6 | 73 | 20 | AAU11552 | Human 5' EST secre |
| 855 | 6 | 0.6 | 62 | 21 | AAU16636 | Bacteriophage 192 | 928 | 6 | 0.6 | 73 | 22 | AAU60392 | Propionibacterium |
| 856 | 6 | 0.6 | 62 | 21 | AAU55685 | Arabidopsis thalia | 929 | 6 | 0.6 | 73 | 22 | ABG26128 | Novel human diagno |
| 857 | 6 | 0.6 | 62 | 21 | AAU59382 | Arabidopsis thalia | 930 | 6 | 0.6 | 73 | 22 | ABE15166 | Human nervous syst |
| 858 | 6 | 0.6 | 62 | 22 | ABE95751 | Human testicular a | 931 | 6 | 0.6 | 73 | 23 | ABG61210 | Tick ixolaris amin |
| 859 | 6 | 0.6 | 62 | 22 | AAU95045 | Human reproductive | 932 | 6 | 0.6 | 74 | 20 | AAU31399 | HEV peptide z12-or |
| 860 | 6 | 0.6 | 62 | 23 | ABG61199 | Tick ixolaris amin | 933 | 6 | 0.6 | 74 | 20 | AAU31398 | HEV peptide z12-or |
| 861 | 6 | 0.6 | 62 | 23 | ABP33541 | Human ORF2514 prot | 934 | 6 | 0.6 | 74 | 22 | AAU45204 | Propionibacterium |
| 862 | 6 | 0.6 | 62 | 23 | ABP10510 | Human ORFX protein | 935 | 6 | 0.6 | 74 | 22 | AAU51521 | Propionibacterium |
| 863 | 6 | 0.6 | 62 | 23 | ABP28116 | Streptococcus poly | 936 | 6 | 0.6 | 74 | 22 | AAE04257 | Human gene 4 enco |
| 864 | 6 | 0.6 | 63 | 17 | ABU95434 | Interleukin-2 BETA | 937 | 6 | 0.6 | 74 | 23 | ABG61211 | Tick ixolaris amin |
| 865 | 6 | 0.6 | 63 | 20 | AAU41340 | Human secreted pro | 938 | 6 | 0.6 | 74 | 23 | ABP34424 | Human ORF3397 prot |
| 866 | 6 | 0.6 | 63 | 21 | AAU59385 | Arabidopsis thalia | 939 | 6 | 0.6 | 74 | 23 | ABP03480 | Human ORFX protein |
| 867 | 6 | 0.6 | 63 | 22 | AAU89874 | C Glutamicum prote | 940 | 6 | 0.6 | 74 | 24 | ABU58181 | Corn stress respon |
| 868 | 6 | 0.6 | 63 | 23 | ABG61200 | Tick ixolaris amin | 941 | 6 | 0.6 | 75 | 19 | AAU62746 | Streptococcus pneu |
| 869 | 6 | 0.6 | 63 | 23 | ABP086619 | Human ORFX protein | 942 | 6 | 0.6 | 75 | 21 | AAU34537 | Arabidopsis thalia |
| 870 | 6 | 0.6 | 63 | 23 | ABP10061 | Human ORFX protein | 943 | 6 | 0.6 | 75 | 21 | AAU55672 | M. pneumoniae ycfB |
| 871 | 6 | 0.6 | 64 | 20 | AAU88786 | Human ORFX protein | 944 | 6 | 0.6 | 75 | 22 | ABG58275 | Human liver peptid |
| 872 | 6 | 0.6 | 64 | 21 | AAU19329 | Polypeptide fragme | 945 | 6 | 0.6 | 75 | 22 | AAU46118 | Propionibacterium |
| 873 | 6 | 0.6 | 64 | 22 | AAU48932 | Arabidopsis thalia | 946 | 6 | 0.6 | 75 | 22 | AAU63762 | Human brain expres |
| 874 | 6 | 0.6 | 64 | 22 | AAU51830 | Propionibacterium | 947 | 6 | 0.6 | 75 | 22 | ABG36683 | Peptide #10720 enc |
| 875 | 6 | 0.6 | 64 | 22 | ABE50622 | Propionibacterium | 948 | 6 | 0.6 | 75 | 23 | ABG61212 | Tick ixolaris amin |
| 876 | 6 | 0.6 | 64 | 22 | AAU82482 | Human immune/haema | 949 | 6 | 0.6 | 76 | 22 | ABG56080 | Human liver peptid |
| 877 | 6 | 0.6 | 64 | 22 | AAU88282 | Human immune/haema | 950 | 6 | 0.6 | 76 | 22 | ABE7244 | Protophilla melanog |
| 878 | 6 | 0.6 | 64 | 23 | AAE28312 | Human Dinp protein | 951 | 6 | 0.6 | 76 | 22 | AAU67291 | Propionibacterium |
| 879 | 6 | 0.6 | 64 | 23 | ABG61201 | Tick ixolaris amin | 952 | 6 | 0.6 | 76 | 22 | ABG08931 | Novel human diagno |
| 880 | 6 | 0.6 | 65 | 6 | AAU50167 | Sequence of the si | 953 | 6 | 0.6 | 76 | 22 | ABE40632 | Peptide #8138 enco |
| 881 | 6 | 0.6 | 65 | 20 | AAU60268 | Human endometrium | 954 | 6 | 0.6 | 76 | 22 | AAU30207 | Novel human secret |
| 882 | 6 | 0.6 | 65 | 21 | AAU52182 | Human secreted pro | 955 | 6 | 0.6 | 76 | 22 | AAU61491 | Human brain expres |
| 883 | 6 | 0.6 | 65 | 21 | AAU55684 | Arabidopsis thalia | 956 | 6 | 0.6 | 76 | 22 | AAU74281 | Human bone marrow |
| 884 | 6 | 0.6 | 65 | 21 | AAU59381 | Arabidopsis thalia | 957 | 6 | 0.6 | 76 | 22 | AAU34393 | Peptide #8430 enco |
| 885 | 6 | 0.6 | 65 | 22 | AAU86070 | Human immune/haema | 958 | 6 | 0.6 | 76 | 23 | ABO8031 | HMM derived consen |

959 6 0.6 76 23 ABG61213
 960 6 0.6 77 19 AAW7751
 961 6 0.6 77 21 AAG56300
 962 6 0.6 77 21 AAG01348
 963 6 0.6 77 22 AAM96471
 964 6 0.6 77 23 ABB08026
 965 6 0.6 77 23 ABG61214
 966 6 0.6 78 22 AAU22026
 967 6 0.6 78 23 ABG61215
 968 6 0.6 78 23 ABA9601
 969 6 0.6 79 20 AAY74049
 970 6 0.6 79 21 AAB34973
 971 6 0.6 79 21 AAG13442
 972 6 0.6 79 22 ABB67938
 973 6 0.6 79 22 ABB70482
 974 6 0.6 79 22 ABG05325
 975 6 0.6 79 22 AAG91567
 976 6 0.6 79 22 AAG91567
 977 6 0.6 79 22 AAU20029
 978 6 0.6 79 23 ABU52000
 979 6 0.6 79 23 ABG61216
 980 6 0.6 79 23 ABA40890
 981 6 0.6 79 23 ABB06851
 982 6 0.6 79 23 ABO03550
 983 6 0.6 80 18 AAW55227
 984 6 0.6 80 20 AAY37446
 985 6 0.6 80 22 AAU44861
 986 6 0.6 80 23 ABP54797
 987 6 0.6 80 23 ABG61217
 988 6 0.6 80 23 ABO03882
 989 6 0.6 80 23 ABO07264
 990 6 0.6 81 17 AAR97308
 991 6 0.6 81 20 AAY35749
 992 6 0.6 81 21 AAG34328
 993 6 0.6 81 21 AAG34383
 994 6 0.6 81 22 AAU54299
 995 6 0.6 81 22 AAM84803
 996 6 0.6 81 22 AAO13375
 997 6 0.6 81 23 ABG61218
 998 6 0.6 81 24 ABR48471
 999 6 0.6 82 19 AAW38476
 1000 6 0.6 82 21 AAG19349

ALIGNMENTS

RESULT 1
 ID AAB62110 standard; Protein; 927 AA.
 XX AC AAB62110;
 XX DT 29-MAY-2001 (first entry)
 XX DE M. bovis Dalton 2d RTX toxin A subunit.
 XX KW Moraxella; antigen; immune response; infection; RTX toxin; vaccine;
 XX KW antibacterial; A subunit.
 XX OS Moraxella bovis.
 XX PN WO200116172-A1.
 XX PD 08-MAR-2001.
 XX PF 31-AUG-2000; 2000WO-AU01048.
 XX PR 31-AUG-1999; 99AU-0002571.
 XX PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
 XX PA (UYME) UNIV MELBOURNE.

PI Farn J, Strugnell R, Tennent J;
 DR WPI; 2001-235092/24.
 DR N-PSDB; AAF57290.
 XX Novel Moraxella bovis antigen useful in compositions for raising immune response in an animal, has protease, lipase or hemolysin activity -
 XX Claim 26; Fig 5; 60pp; English.
 PS The invention relates to new Moraxella bovis antigens and nucleic acid sequences encoding these antigenic polypeptides. The antigenic polypeptides and polynucleotides are useful for raising an immune response in an animal directed against Moraxella, preferably against M. bovis or M. catarrhalis, and for treating Moraxella infections. The present sequence represents the amino acid sequence of the A subunit of the RTX toxin from M. bovis Dalton 2d.
 XX SQ Sequence 927 AA;
 Query Match 100.0%; Score 927; DB 22; Length 927;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 927; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSNIVIKSNIQAGLNSTKSGLNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
 DB 1 MSNIVIKSNIQAGLNSTKSGLNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
 QY 61 NHTETAKSVDTVQNFSLTQTGTIAISATKLEKFLQKHSNKLAKGLSDSVENIDRLKGLKA 120
 DB 61 NHTETAKSVDTVQNFSLTQTGTIAISATKLEKFLQKHSNKLAKGLSDSVENIDRLKGLKA 120
 QY 121 SNVLSTLSPFLGALAGIELDSLIIKGGDAAPDALAKASIDLINEITGNLSQSTQTIEAFS 180
 DB 121 SNVLSTLSPFLGALAGIELDSLIIKGGDAAPDALAKASIDLINEITGNLSQSTQTIEAFS 180
 QY 181 SOLAKLGSTTISQAKGFSNIGNKLNLFNFSKTNLGLLEITIGLLSGISAGFALADKNASTGK 240
 DB 181 SOLAKLGSTTISQAKGFSNIGNKLNLFNFSKTNLGLLEITIGLLSGISAGFALADKNASTGK 240
 QY 241 KVAAGFELSNOVITGNVTKAISSVYLAQRVAAGLSTTGAVAAALITSSIMLAISPLAFNAA 300
 DB 241 KVAAGFELSNOVITGNVTKAISSVYLAQRVAAGLSTTGAVAAALITSSIMLAISPLAFNAA 300
 QY 301 DKFNHANALDEFKQFRKFGYDGDHLLAEYQVGVGTIEASLTITISTALGAVSAGVSAAV 360
 DB 301 DKFNHANALDEFKQFRKFGYDGDHLLAEYQVGVGTIEASLTITISTALGAVSAGVSAAV 360
 QY 361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLOKILEWEKQNGQNYFDKGYD 420
 DB 361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLOKILEWEKQNGQNYFDKGYD 420
 QY 421 SRYAAVLANNKFLSELNKELEAEVIAITQORWNNIGELAGITKLGERIKSGKAYADA 480
 DB 421 SRYAAVLANNKFLSELNKELEAEVIAITQORWNNIGELAGITKLGERIKSGKAYADA 480
 QY 481 FEDGKVEAGSNITLDAKTGIIIDISNNGKKTQALHFTSPLLTAGTESRRLTNGKYSYI 540
 DB 481 FEDGKVEAGSNITLDAKTGIIIDISNNGKKTQALHFTSPLLTAGTESRRLTNGKYSYI 540
 QY 541 NKLKFRGRVKNVQVTDGEASKLDIFSIVQVASTEGTDEIGLIVNAKAGNDDIFVGGGKM 600
 DB 541 NKLKFRGRVKNVQVTDGEASKLDIFSIVQVASTEGTDEIGLIVNAKAGNDDIFVGGGKM 600
 QY 601 NIDGGGHDHVRVYSGKGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKQETKVGK 660
 DB 601 NIDGGGHDHVRVYSGKGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKQETKVGK 660
 QY 661 RTETIQVRYELRKVGVGQSTDNLSKVEEVIGSQFNDFPKGSKFNDFIHSGGGDDLLDG 720
 DB 661 RTETIQVRYELRKVGVGQSTDNLSKVEEVIGSQFNDFPKGSKFNDFIHSGGGDDLLDG 720
 QY 721 GAGDDRLFGGKGNDRLSGDGDDLLDGGSGDDVINGAGNDVYIFRKGNDTLTYDGTGN 780

```
Db 721 GAGDRLFGKGNDRLSGDEGDLDDGGSGDDVLNGAGNDVYFRKGDGNDLYDGTGN 780
Qy 781 DKLAFADANISDIMERKTEGIIIVKRNHSGSINIPRWYITSNLQYQSNKTDHKIEOLI 840
Db 781 DKLAFADANISDIMERKTEGIIIVKRNHSGSINIPRWYITSNLQYQSNKTDHKIEOLI 840
Qy 841 GKDGSYITSDIQDKILQDKDGTVITSQELKKLADENKSKLSASDIASSLNKLVGSMAL 900
Db 841 GKDGSYITSDIQDKILQDKDGTVITSQELKKLADENKSKLSASDIASSLNKLVGSMAL 900
Qy 901 FGTSNVSSNALQPIPTQPTQGIILAPSV 927
Db 901 FGTSNVSSNALQPIPTQPTQGIILAPSV 927

RESULT 2
AAR86998
ID AAR86998 standard; Protein; 758 AA.
XX
AC AAR86998;
XX
DT 04-JUL-1996 (first entry)
XX
DE Enterohaemorrhagic E.coli hlyA gene product.
XX
KW Enterohaemorrhagic Escherichia coli; virulent; EHEC; O157:H7 serotype;
KW detection; probe; primer; hlyA gene; enterohaemorrhagic colitis;
KW haemolytic uremic syndrome; mesenteric adenitis.
XX
OS Escherichia coli (enterohaemorrhagic).
XX
PN US5475098-A.
XX
PD 12-DEC-1995.
XX
PF 14-JUN-1994; 94US-0258188.
XX
PR 14-JUN-1994; 94US-0258188.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Hall RH, Xu JG;
XX
XX WPI; 1996-048546/05.
XX N-PSDB; AAT08098.
XX
PT Enterohaemorrhagic E. coli (EHEC) nucleic acid sequences - useful
PT for probe and primer design for sensitive and specific detection of
PT EHEC
XX
PS Claim 1; Columns 37-42; 32pp; English.
XX
CC Enterohaemorrhagic E.coli (EHEC) associated with enterohaemorrhagic
CC colitis, haemolytic uremic syndrome and mesenteric adenitis have
CC been found to carry a hlyA gene and a hlyB gene, separated by an
CC intergenic region. The hly genes and the intergenic region are
CC absent from bacteria not associated with these diseases and so
CC provide a useful target for detecting EHEC pathogens, esp. O157:H7
CC serotype E.coli. The present sequence is that of the protein
CC encoded by the EHEC hlyA gene.
XX
SQ Sequence 758 AA;
Query Match 1.5%; Score 14; DB 17; Length 758;
Best Local Similarity 100.0%; Pred. No. 0.00048;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 379 ISGILEASKQAMFE 392
Db 152 ISGILEASKQAMFE 165
```

```
RESULT 3
AAW22159
ID AAW22159 standard; Protein; 1049 AA.
XX
AC AAW22159;
XX
DT 16-FEB-1998 (first entry)
XX
DE ApxIIIB protein.
XX
KW RTX toxin; apxICA gene; apxIBD gene; apxIIAB'C gene; apxIIABCD gene;
KW repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
KW therapy; A. pleuropneumoniae infection; swine pleuropneumonia.
XX
OS Actinobacillus pleuropneumoniae.
XX
PN CA2170839-A.
XX
PD 02-SEP-1996.
XX
PF 01-MAR-1996; 96CA-2170839.
XX
PR 01-MAR-1995; 95US-0396244.
XX
PA (UYGU-) UNIV GUELPH.
XX
PI MacInnes J, Mallard B, Ricciatti P, Rosendal S;
XX
XX WPI; 1997-245536/23.
XX N-PSDB; AAT73220.
XX
PT Preparations of microorganisms producing cell-associated RTX toxins
PT - especially for production of vaccines against swine
PT pleuro-pneumonia
XX
PS Disclosure; Pages 107-110; 151pp; English.
XX
XX AAW22151-W22161 represent A. pleuropneumoniae RTX (repeat in toxins)
XX toxins. These sequence are encoded by the apxICA, apxIBD, apxIIAB'C, and
XX apxIIABCD genes (see AAT73217-T73220), and can be expressed by
XX microorganisms used in the preparations of the invention. The
XX preparations are bacterial preparations comprising one or more isolated
XX and purified strains of a microorganism that produces one or more RTX
XX toxins, where the strains have at least one cell-associated RTX toxin.
XX The preparations are used for production of vaccines for the prophylaxis
XX and treatment of infectious diseases caused by microorganisms that
XX produce RTX toxins, where the strains have been attenuated or
XX inactivated. The vaccines are preferably against Actinobacillus
XX pleuropneumoniae infection (swine pleuropneumonia). It has been found
XX that A. pleuropneumonia produces significant quantities of
XX cell-associated RTX toxins when cultured under certain conditions, and
XX that the whole-cell protein composition of the cultures corresponds to
XX necropsy from the pleural fluid of infected swine. Vaccination with a
XX bacterin prepared from heat-inactivated cultures having significant
XX quantities of cell-associated RTX toxins give significant protection of
XX swine against challenge with homologous strains.
XX
SQ Sequence 1049 AA;
Query Match 1.5%; Score 14; DB 18; Length 1049;
Best Local Similarity 100.0%; Pred. No. 0.00064;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 KVAAGFELSNQVIG 254
Db 275 KVAAGFELSNQVIG 288

RESULT 4
AAV51412
ID AAV51412 standard; protein; 1049 AA.
XX
```


AC AAY51412;
 XX
 DT 05-MAY-2000 (first entry)
 XX
 DE
 XX
 DE A. pleuropneumoniae apxIIIA protein.
 XX
 KW RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
 KW antiinflammatory; antiarthritic; antiabortive; treatment; pneumonia;
 KW pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
 KW shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
 KW urinary infection; peritonitis; meningitis; gastroenteritis;
 KW passive immunization; apxIIIA.
 XX
 OS Actinobacillus pleuropneumoniae.
 XX
 FN US6019984-A.
 XX
 PD 01-FEB-2000.
 XX
 XX 23-DEC-1996; 96US-0772270.
 XX
 XX 01-MAR-1995; 95US-0396244.
 XX
 PA (UYGU-) UNIV GUELPH.
 XX
 PI Mallard B, Rosendal S, MacInnes J, Ricciatti P.
 XX
 XX WPI; 2000-146864/13.
 DR
 DR N-PSDB; AAZ88587.
 XX
 XX Bacterial preparation comprising microorganisms which produce a member
 PT of the Repeats in Toxins (RTX) family, useful for treating swine
 PT pleuropneumonia, arthritis in swine, shipping fever and abortion in
 PT cattle, and sleepy foal disease -
 XX
 XX Disclosure; Column 77-84; 96pp; English.
 PS
 CC This invention describes a novel bacterial preparation (I) which
 CC comprises one or more isolated and purified strain(s) of a microorganism,
 CC cultured in tryptone yeast extract (TYE) broth, which produces one or
 CC more RTX toxins (belonging to the family of toxins referred to as Repeats
 CC in Toxins), where the strain(s) have at least one RTX toxin which is
 CC cell-associated. The products of the invention have immunostimulatory,
 CC antimicrobial, antiinflammatory, antiarthritic and antiabortive activity.
 CC The bacterial preparation may be used as vaccines for the prophylaxis and
 CC treatment of infectious diseases caused by strains of microorganisms
 CC which produce one or more RTX toxins. The infectious diseases are swine
 CC pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
 CC septicemia, nephritis, endocarditis and arthritis in piglets; shipping
 CC fever and abortion in cattle; whooping cough, sleepy foal disease or
 CC joint ill (purulent nephritis, arthritis) in foals; septicemia,
 CC polyarthritis and abortion in horses; and urinary infections,
 CC peritonitis, meningitis, and gastroenteritis. The bacterial preparations
 CC may also be used to prepare antibodies which may be used as a means of
 CC passive immunization. This sequence represents the Actinobacillus
 CC pleuropneumoniae apxIIIA protein described in the method of the
 CC invention.
 XX
 SQ Sequence 1049 AA;
 XX
 Query Match 1.5%; Score 14; DB 21; Length 1049;
 Best Local Similarity 100.0%; Pred. No. 0.00064;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 241 KVAAGFELSNOVIG 254
 |||||
 Db 275 KVAAGFELSNOVIG 288
 |||||
 RESULT 5
 AAR54781
 ID AAR54781 standard; Protein; 1244 AA.
 XX

AC AAR54781;
 XX
 DT 25-MAR-2003 (updated)
 DT 15-OCT-1994 (first entry)
 XX
 DE Leukotoxin ApxIIIA.
 XX
 KW Leukotoxin; ApxIIIA; pleuropneumonia; vaccine; diagnostic.
 XX
 OS Actinobacillus pleuropneumoniae.
 XX
 FH Key Location/Qualifiers
 FT Domain 327..345 /label= transmembrane domain
 FT Domain 488..503 /label= transmembrane domain
 FT Domain 570..587 /label= transmembrane domain
 FT
 XX WO9409821-A1.
 XX
 PD 11-MAY-1994.
 XX
 XX 04-NOV-1993; 93WO-US10500.
 XX
 XX 05-NOV-1992; 92US-0972229.
 PR 03-JUN-1993; 93US-0072285.
 XX
 PA (CORR) CORNELL RES FOUND INC.
 XX
 XX Chang Y;
 XX
 DR WPI; 1994-167130/20.
 DR N-PSDB; AAQ64827.
 XX
 PT DNA encoding Actinobacillus pleuropneumoniae leukotoxin - used to
 PT prepare prods. for use in vaccines for porcine pleuropneumonia
 PT and as diagnostic reagents
 XX
 PS Disclosure; Fig 3; 65pp; English.
 CC
 CC The ApxIIIA gene having the sequence given in AAQ64827 was isolated
 CC from phage lambda clones yfc 26-28 and yfc 31-32 of a genomic
 CC library of A. pleuropneumoniae ser. 2. The gene encodes a novel
 CC leukotoxin (sequence AAR54781).
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 1244 AA;
 XX
 Query Match 1.5%; Score 14; DB 15; Length 1244;
 Best Local Similarity 100.0%; Pred. No. 0.00075;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 241 KVAAGFELSNOVIG 254
 |||||
 Db 448 KVAAGFELSNOVIG 461
 |||||
 RESULT 6
 AAR12561
 ID AAR12561 standard; Protein; 956 AA.
 XX
 AC AAR12561;
 XX
 DT 25-MAR-2003 (updated)
 DT 26-JUL-1991 (first entry)
 XX
 DE APPA haemolysin antigen encoded by appA.
 XX
 KW Vaccine; porcine pleuropneumonia; haemolysin; cytotoxin.
 XX
 OS Actinobacillus pleuropneumoniae serotype 5.
 XX

PN W09106653-A.
 XX 16-MAY-1991.
 XX 31-OCT-1990; 90WO-US06350.
 XX 31-OCT-1989; 89US-0429273.
 XX (TEXA) UNIV TEXAS A & M SYSTEM.
 XX Struck DK, Young RF, Chang YF;
 XX WPI; 1991-164202/22.
 XX N-PSDB; AAQ11873.
 XX DNA encoding Actinobacillus pleuropneumoniae haemolysin - used
 PT for producing haemolysin antigen for vaccinating pigs against
 PT porcine pleuro-pneumonia.
 XX Claim 8; Fig 1; 28pp; English.
 XX The sequence is encoded by the appA gene carried on a 3.8 kb
 CC section of clone yfc5 which was isolated by immunological screen-
 CC ing. It forms the appCA antigen with the appC protein encoded on
 CC the same section of DNA. The appA antigen is not in itself
 CC haemolytic but does elicit an immune response and can be used in
 CC vaccines.
 CC See also AAR12127.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX Sequence 956 AA;
 SQ Query Match 1.2%; Score 11; DB 12; Length 956;
 Best Local Similarity 100.0%; Pred. No. 0.43;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 369 ALLVAGVTGLI 379
 DB 392 ALLVAGVTGLI 402
 RESULT 7
 ID AAW22156
 XX AAW22156 standard; Protein; 956 AA.
 AC AAW22156;
 XX 16-FEB-1998 (first entry)
 DE ApxIIC protein.
 XX RTX toxin; apxICA gene; apxIBD gene; apxIIAB/C gene; apxIIABCD gene;
 KW repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
 KW therapy; A. pleuropneumoniae infection; swine pleuropneumonia.
 XX Actinobacillus pleuropneumoniae.
 OS CA2170839-A.
 XX 02-SEP-1996.
 XX 01-MAR-1996; 96CA-2170839.
 XX 01-MAR-1995; 95US-0396244.
 XX (UYGU-) UNIV GUELPH.
 XX MacInnes J, Mallard B, Ricciatti P, Rosendal S;
 XX WPI; 1997-245536/23.
 DR N-PSDB; AAT73219.
 XX Preparations of microorganisms producing cell-associated RTX toxins

PT - especially for production of vaccines against swine
 PT pleuro-pneumonia
 XX Disclosure; Pages 101-104; 151pp; English.
 XX AAW22151-W22161 represent A. pleuropneumoniae RTX (repeat in toxins)
 CC toxins. These sequence are encoded by the apxICA, apxIBD, apxIIAB/C, and
 CC apxIIABCD genes (see AAT73217-T73220), and can be expressed by
 CC microorganisms used in the preparations of the invention. The
 CC preparations are bacterial preparations comprising one or more isolated
 CC and purified strains of a microorganism that produces one or more RTX
 CC toxins, where the strains have at least one cell-associated RTX toxin.
 CC The preparations are used for production of vaccines for the prophylaxis
 CC and treatment of infectious diseases caused by microorganisms that
 CC produce RTX toxins, where the strains have been attenuated or
 CC inactivated. The vaccines are preferably against Actinobacillus
 CC pleuropneumoniae infection (swine pleuropneumonia). It has been found
 CC that A. pleuropneumonia produces significant quantities of
 CC cell-associated RTX toxins when cultured under certain conditions, and
 CC that the whole-cell protein composition of the cultures corresponds to
 CC the whole-cell protein profiles obtained from cells recovered at
 CC necropsy from the pleural fluid of infected swine. Vaccination with a
 CC bacterin prepared from heat-inactivated cultures having significant
 CC quantities of cell-associated RTX toxins give significant protection of
 CC swine against challenge with homologous strains.
 XX Sequence 956 AA;
 SQ Query Match 1.2%; Score 11; DB 18; Length 956;
 Best Local Similarity 100.0%; Pred. No. 0.43;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 369 ALLVAGVTGLI 379
 DB 392 ALLVAGVTGLI 402
 RESULT 8
 ID AAY51410
 XX AAY51410 standard; protein; 956 AA.
 AC AAY51410;
 XX 05-MAY-2000 (first entry)
 DE A. pleuropneumoniae clyIIA protein.
 XX RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
 KW antinflammatory; antiarthritic; antiabortive; treatment; pneumonia;
 KW pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
 KW shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
 KW urinary infection; peritonitis; meningitis; gastroenteritis;
 KW passive immunization; clyIIA.
 XX Actinobacillus pleuropneumoniae.
 OS US6019984-A.
 XX 01-FEB-2000.
 XX 23-DEC-1996; 96US-0772270.
 XX 01-MAR-1995; 95US-0396244.
 XX (UYGU-) UNIV GUELPH.
 XX Mallard B, Rosendal S, MacInnes J, Ricciatti P;
 XX WPI; 2000-146864/13.
 DR N-PSDB; AA288586.
 XX Bacterial preparation comprising microorganisms which produce a member
 PT of the Repeats in Toxins (RTX) family, useful for treating swine

PT pleuropneumonia, arthritis in swine, shipping fever and abortion in
 PT cattle, and sleepy foal disease -
 XX Disclosure; Column 67-72; 96pp; English.
 XX
 CC This invention describes a novel bacterial preparation (I) which
 CC comprises one or more isolated and purified strain(s) of a microorganism,
 CC cultured in tryptone yeast extract (TYE) broth, which produces one or
 CC more RTX toxins (belonging to the family of toxins referred to as Repeats
 CC in Toxins), where the strain(s) have at least one RTX toxin which is
 CC cell-associated. The products of the invention have immunostimulatory,
 CC antimicrobial, antiinflammatory, antiarthritic and antiabortive activity.
 CC The bacterial preparation may be used as vaccines for the prophylaxis and
 CC treatment of infectious diseases caused by strains of microorganisms
 CC which produce one or more RTX toxins. The infectious diseases are swine
 CC pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
 CC septicemia, nephritis, endocarditis and arthritis in piglets; shipping
 CC fever and abortion in cattle; whooping cough, sleepy foal disease or
 CC joint ill (purulent nephritis, arthritis) in foals; septicemia,
 CC polyarthritis and abortion in horses; and urinary infections,
 CC peritonitis, meningitis, and gastroenteritis. The bacterial preparations
 CC may also be used to prepare antibodies which may be used as a means of
 CC passive immunization. This sequence represents the Actinobacillus
 CC pleuropneumoniae clyIIA protein described in the method of the invention.
 XX
 SQ Sequence 956 AA;

Query Match 1.2%; Score 11; DB 21; Length 956;
 Best Local Similarity 100.0%; Pred. No. 0.43;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 369 ALLVAGVTGLI 379
 Db 392 ALLVAGVTGLI 402

RESULT 9
 AAW22152
 ID AAW22152 standard; Protein; 1022 AA.

XX AAW22152;

DT 16-FEB-1998 (first entry)

DE ApxIA protein.

KW RTX toxin; apxICA gene; apxIBD gene; apxIIAB'C gene; apxIIABCD gene;
 KW repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
 KW therapy; A. pleuropneumoniae infection; swine pleuropneumonia.

XX Actinobacillus pleuropneumoniae.

XX CA2170839-A.

XX 02-SEP-1996.

XX 01-MAR-1996; 96CA-2170839.

XX 01-MAR-1995; 95US-0396244.

XX (UYGU-) UNIV GUELPH.

XX MacInnes J, Mallard B, Ricciatti P, Rosendal S;

XX WPI; 1997-245536/23.

XX N-PSDB; AAT73217.

XX Preparations of microorganisms producing cell-associated RTX toxins
 PT - especially for production of vaccines against swine
 PT pleuro-pneumonia

XX Disclosure; Pages 88-91; 151pp; English.

CC AAW22151-W22161 represent A. pleuropneumoniae RTX (repeat in toxins)
 CC toxins. These sequence are encoded by the apxICA, apxIBD, apxIIAB'C, and
 CC apxIIABCD genes (see AAT73217-T73220), and can be expressed by
 CC microorganisms used in the preparations of the invention. The
 CC preparations are bacterial preparations comprising one or more isolated
 CC and purified strains of a microorganism that produces one or more RTX
 CC toxins, where the strains have at least one cell-associated RTX toxin.
 CC The preparations are used for production of vaccines for the prophylaxis
 CC and treatment of infectious diseases caused by microorganisms that
 CC produce RTX toxins, where the strains have been attenuated or
 CC inactivated. The vaccines are preferably against Actinobacillus
 CC pleuropneumoniae infection (swine pleuropneumonia). It has been found
 CC that A. pleuropneumoniae produces significant quantities of
 CC cell-associated RTX toxins when cultured under certain conditions, and
 CC that the whole-cell protein composition of the cultures corresponds to
 CC the whole-cell protein profiles obtained from cells recovered at
 CC necropsy from the pleural fluid of infected swine. Vaccination with a
 CC bacterin prepared from heat-inactivated cultures having significant
 CC quantities of cell-associated RTX toxins give significant protection of
 CC swine against challenge with homologous strains.

XX SQ Sequence 1022 AA;

Query Match 1.2%; Score 11; DB 18; Length 1022;

Best Local Similarity 100.0%; Pred. No. 0.45;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 266 AQRVAAGLSTT 276

Db 289 AQRVAAGLSTT 299

RESULT 10

AAW51406

ID RAYS1406 standard; protein; 1022 AA.

XX AAW51406;

XX 05-MAY-2000 (first entry)

XX A. pleuropneumoniae apxIA protein.

DE RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
 DE antiinflammatory; antiarthritic; antiabortive; treatment; pneumonia;
 KW pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
 KW shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
 KW urinary infection; peritonitis; meningitis; gastroenteritis;
 KW passive immunization; apxIA.

XX Actinobacillus pleuropneumoniae.

XX US6019984-A.

XX 01-FEB-2000.

XX 23-DEC-1996; 96US-0772270.

XX 01-MAR-1995; 95US-0396244.

XX (UYGU-) UNIV GUELPH.

XX Mallard B, Rosendal S, MacInnes J, Ricciatti P;

XX WPI; 2000-146864/13.

XX N-PSDB; AAZ88584.

XX Bacterial preparation comprising microorganisms which produce a member
 PT of the Repeats in Toxins (RTX) family, useful for treating swine
 PT pleuropneumonia, arthritis in swine, shipping fever and abortion in
 PT cattle, and sleepy foal disease -

XX Disclosure; Column 47-52; 96pp; English.


```

Db      11 AERVIAITQQ 20

RESULT 13
AAV33929
ID      AAV33929 standard; Protein; 490 AA.
XX
XX
AC      AAV33929;
XX
DT      09-NOV-1999 (first entry)
XX
DE      Leukotoxin carrier protein (LKT 114).
XX
KW      Myostatin; mouse; rabbit; human; baboon; bovine; porcine; ovine; chick;
KW      turkey; zebrafish; immune response; vaccine; body weight; muscle mass;
KW      mammary gland tissue; lactation; feed uptake; muscle degeneration;
KW      GDF11 activity; LKT 114; leukotoxin carrier protein.
XX
OS      Pasteurella haemolytica.
XX
XX      WO9942573-A1.
XX
PN      26-AUG-1999.
XX
PD      19-FEB-1999; 99WO-CA00128.
XX
PF      19-FEB-1998; 98US-0075213.
XX
PR      (BIOS-) BIOSTAR INC.
XX
PA      Barker CA, Morsey M;
XX
PI      WPI; 1999-527471/44.
XX
DR      N-PSDB; AAX99361.
XX
PT      New myostatin peptide, multimers and immunoconjugates for eliciting
PT      an immune response in a vertebrate against a myostatin immunogen
XX
XX      Example 2; Fig 15; 109pp; English.
XX
CC      The invention provides myostatin peptides consisting of 3-100 amino
CC      acids, derived from a region of mouse, rabbit, human, baboon, bovine,
CC      porcine, ovine, chick, turkey or zebrafish myostatin (see sequences
CC      AAV3390-939). The myostatin peptides are derived preferably from a
CC      region of amino acid residues 1-275, 25-300, 50-325 or 75-350 of the
CC      above sequences. The peptides and the nucleic acids encoding the peptides
CC      are useful as vaccines for eliciting an immune response in a vertebrate
CC      against a myostatin immunogen. They result in increasing body weight,
CC      muscle mass, number and size of muscle cells, muscle strength, mammary
CC      gland tissue, lactation, appetite or feed uptake, life span of the
CC      vertebrate, and cause a reduction in body fat content, useful for muscle
CC      wasting conditions. The vaccines are also useful for treating a disorder
CC      which comprises degeneration or wasting of muscle in a vertebrate, and
CC      useful for modulating GDF11 activity. The present sequence represents a
CC      leukotoxin carrier protein (LKT 114) which is used as an immunological
CC      carrier protein to form a myostatin immunoconjugate.
XX
SQ      Sequence 490 AA;
Query Match 1.1%; Score 10; DB 20; Length 490;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      265 LAQRVAAGLS 274
Db      256 LAQRVAAGLS 265
|||||
RESULT 14
AAW03943
ID      AAW03943 standard; Protein; 544 AA.
XX
XX      AAW03943;
AC

XX      20-NOV-1996 (first entry)
XX
XX      LKT-GnRH protein fusion from pCB111.
XX
KW      Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
KW      fusion protein; immunogen; vaccine; fertility control;
KW      contraceptive; sterilisation.
XX
OS      Chimeric Pasteurella haemolytica A1 strain B122;
OS      Chimeric synthetic.
XX
XX      Key Location/Qualifiers
XX      Domain 1..493
XX      Domain /label= LKT
XX      Domain /label= GnRH_repeat_domain
XX
PN      WO9624675-A1.
XX
XX      15-AUG-1996.
XX
PD      24-JAN-1996; 96WO-CA00049.
XX
PF      10-FEB-1995; 95US-0387156.
XX
PR      (UYSA-) UNIV SASKATCHEWAN.
XX
PA      Manns JG, Potter AA;
XX
PI      WPI; 1996-384447/38.
XX
DR      N-PSDB; AAT37177.
XX
PT      Gonadotropin-releasing hormone multimer fusion proteins - with
PT      leukotoxin polypeptide for increased immunogenicity, useful in
PT      antifertility vaccine prodn.
XX
XX      Claim 8; Fig 7A-7E; 87pp; English.
XX
CC      A chimeric protein (AAW03943) is composed of a fusion between
CC      a truncated leukotoxin (LKT-111) from Pasteurella haemolytica
CC      and a 4-copy gonadoliberin-releasing hormone (GnRH) repeat
CC      sequence (see also AAW03944). It is the product of a chimeric
CC      gene (AAT37177) produced by deleting an approx. 1300 bp sequence
CC      from pCB113 (see also AAT37176) coding for amino acids 352-784
CC      of LKT-352. Recombinant plasmid pCB111 (LKT 111:4 copy GnRH,
CC      ATCC 69748) was obtd. Escherichia coli transformants produced
CC      the chimeric protein, which is useful as a vaccine for fertility
CC      control, esp. immunological sterilisation of domestic or farm
CC      animals.
XX
SQ      Sequence 544 AA;
Query Match 1.1%; Score 10; DB 17; Length 544;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      265 LAQRVAAGLS 274
Db      256 LAQRVAAGLS 265
|||||
RESULT 15
AAW79570
ID      AAW79570 standard; Protein; 544 AA.
XX
XX      AAW79570;
AC

XX      24-DEC-1998 (first entry)
XX
XX      LKT-GnRH chimeric protein.
XX
XX      Chimera; pCB111; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;
KW

```

KW cytotoxic activity; antigen presentation; immune response; vaccine;
 KW tumour.
 XX Synthetic.
 OS WO9806848-A1.
 PN 19-FEB-1998.
 XX 08-AUG-1997; 97WO-CA00559.
 XX 09-AUG-1996; 96US-0694865.
 XX (UYSA-) UNIV SASKATCHEWAN.
 XX Manns JG, Potter AA;
 XX WPI; 1998-159540/14.
 DR N-PSDB; AAV61532.
 XX Chimeric protein of leukotoxin and gonadotropin releasing hormone -
 PT useful for, e.g. preparation of vaccines for reduction of incidence
 PT of mammary tumours in mammals
 XX Disclosure; Figure 7.1-5; 118pp; English.
 XX The present sequence represents the LKT-GnRH chimeric protein from
 CC pCB111. This plasmid contains the LKT 111 polypeptide fused to
 CC four copies of the GnRH peptide. This chimera lacks cytotoxic activity
 CC which enables there to be an increase in antigen presentation and thus an
 CC optimal immune response. The removal of this region also enables the
 CC truncated LKT to be expressed at much higher levels and allows the amount
 CC of antigen administered to be reduced. This chimeric protein comprises a
 CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The
 CC chimeric protein can be used as a vaccine to help reduce the incidence of
 CC mammary tumours in a mammalian individual.
 XX SQ Sequence 544 AA;
 Query Match 1.1%; Score 10; DB 19; Length 544;
 Best Local Similarity 100.0%; Pred. NO. 2.3;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 265 LAQRVAAGLS 274
 DB 256 LAQRVAAGLS 265
 RESULT 16
 AAE04636
 ID AAE04636 standard; Protein; 608 AA.
 XX AAE04636;
 XX 10-SEP-2001 (first entry)
 XX Pasteurella haemolytica modified leukotoxin 66 (lkt66) protein.
 DE Leukotoxin 66; lkt66; respiratory disease; infection; therapy;
 KW immunostimulant; antibacterial; vaccine; transgenic plant;
 KW animal feed; mutant; mutein.
 XX Pasteurella haemolytica.
 OS Synthetic.
 XX WO200144289-A2.
 XX 21-JUN-2001.
 XX 15-DEC-2000; 2000WO-CA01498.
 XX 17-DEC-1999; 99US-0172148.

PA (UYGU-) UNIV GUELPH OFFICE.
 XX Lo RYC, Shewen PE, Lee RWH, Hodgins D, Strommer JN;
 XX WPI; 2001-408470/43.
 DR N-PSDB; AAD08975.
 XX Modified leukotoxin polypeptide is useful in a vaccine to prevent or
 PT treat Mannheimia (Pasteurella) infection (particularly M. haemolytica
 PT infection), and disease associated with a leukotoxin, e.g., respiratory
 PT disease -
 XX Claim 2; Fig 2; 70pp; English.
 XX The present sequence is Pasteurella (Mannheimia) haemolytica
 CC modified leukotoxin-66 (lkt66) protein. The modification comprises
 CC the removal of amino acids within the hydrophobic transmembrane
 CC domain of a full length leukotoxin protein. Modified leukotoxin
 CC sequences are used in vaccines to treat or prevent diseases associated
 CC with leukotoxin, e.g., respiratory disease, and Mannheimia infection
 CC (particularly M. haemolytica infection). In addition, the vaccine
 CC is used to prepare a medicament. Furthermore, the plant transformed
 CC with modified leukotoxin sequences is fed to an animal such as a
 CC ruminant, to prevent or treat respiratory diseases.
 XX SQ Sequence 608 AA;
 Query Match 1.1%; Score 10; DB 22; Length 608;
 Best Local Similarity 100.0%; Pred. NO. 2.5;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 443 AERVAITQO 452
 DB 116 AERVAITQO 125
 RESULT 17
 AAW79573
 ID AAW79573 standard; Protein; 695 AA.
 XX AAW79573;
 AC 24-DEC-1998 (first entry)
 XX LKT-GnRH chimeric protein.
 DE Chimera; pCB122; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;
 KW cytotoxic activity; antigen presentation; immune response; vaccine;
 KW tumour.
 XX Synthetic.
 XX WO9806848-A1.
 XX 19-FEB-1998.
 XX 08-AUG-1997; 97WO-CA00559.
 XX 09-AUG-1996; 96US-0694865.
 XX (UYSA-) UNIV SASKATCHEWAN.
 XX Manns JG, Potter AA;
 XX WPI; 1998-159540/14.
 DR N-PSDB; AAV61535.
 XX Chimeric protein of leukotoxin and gonadotropin releasing hormone -
 PT useful for, e.g. preparation of vaccines for reduction of incidence
 PT of mammary tumours in mammals
 XX Claim 9; Figure 9.1-6; 118pp; English.

CC The present sequence represents the LKT-GnRH chimeric protein from
 CC pCB122. This plasmid contains the LKT 111 polypeptide fused to sixteen
 CC copies of the GnRH peptide, in the pattern of: 8 copies of GnRH-LKT 111-8
 CC copies of GnRH. This chimera lacks cytotoxic activity which enables
 CC there to be an increase in antigen presentation and thus an optimal
 CC immune response. The removal of this region also enables the truncated
 CC LKT to be expressed at much higher levels and allows the amount of
 CC antigen administered to be reduced. This chimeric protein comprises a
 CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The
 CC chimeric protein can be used as a vaccine to help reduce the incidence of
 CC mammary tumours in a mammalian individual.

XX Sequence 695 AA;

Query Match 1.1%; Score 10; DB 19; Length 695;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
 |||||
 Db 359 LAQRVAAGLS 368

RESULT 18
 AAY58361
 ID AAY58361 standard; Protein; 695 AA.

XX AC AAY58361;

XX DT 27-MAR-2000 (first entry)

XX DE Leukotoxin/gonadotropin-releasing hormone fusion protein.

XX KW Leukotoxin; gonadotropin-releasing hormone; GnRH; immunosterilisation;
 KW immunoconception; vaccine; feline; canine; equine; cervine.

XX OS Chimeric - Pasteurella haemolytica.
 OS Chimeric - Mammalia.

PN WO9962545-A2.

PD 09-DEC-1999.

PF 28-MAY-1999; 99WO-CA00493.

PR 04-JUN-1998; 98US-0088024.

PR 06-MAY-1999; 99US-0306689.

PA (BIOS-) BIOSTAR INC.

PI Robbins SC;

XX WPI; 2000-086857/07.

DR N-PSDB; AAZ55700.

PT Hormone immunogens, analogues or antibodies used to manufacture
 PT vaccines for suppression of reproductive behavior and fertility in
 PT vertebrates -

XX Claim 20; Fig 6A-6F; 88pp; English.

XX This sequence represents a fusion protein comprising gonadotropin-
 CC releasing hormone (GnRH) immunogens and a Pasteurella
 CC haemolytica leukotoxin (LKT) protein. The fusion protein comprises, in
 CC the N to C-terminal direction, a synthetic peptide sequence (AAY58364),
 CC an eight copy GnRH multimer (composed of two copies of the 4xGnRH
 CC multimer sequence of AAY58363), the LKT protein (which functions as a
 CC carrier protein), and a second eight copy GnRH multimer. The fusion
 CC protein may be used in a vaccine composition for prepubertal
 CC administration to a vertebrate subject to result in prolonged suppression
 CC of reproductive behaviour and/or fertility. GnRH immunogens, analogues or
 CC antibodies that cross-react with endogenous GnRH of a vertebrate subject
 CC are used to manufacture a composition or vaccine for immunosterilisation

CC or immunoconception of feline, canine, equine or cervine subjects.
 CC The vaccines are used to suppress reproductive behaviour and/or
 CC fertility for at least 10 months. The prepubertal administration
 CC results in a prolonged, long-term suppression of testicular development
 CC and/or function in males, or a prolonged, long-term suppression of
 CC ovarian development and/or function in females. The methods provide a
 CC viable and desirable alternative to surgical forms of sterilisation that
 CC are currently used.

XX Sequence 695 AA;

Query Match 1.1%; Score 10; DB 21; Length 695;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
 |||||
 Db 359 LAQRVAAGLS 368

RESULT 19

AAY58133

ID AAY58133 standard; Protein; 695 AA.

XX AC AAY58133;

XX DT 07-MAR-2000 (first entry)

XX DE Gonadotropin releasing hormone-leukotoxin fusion protein.

XX KW Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;
 KW antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;
 KW non-androgenic; steroid; reduction; weight gain; muscle distribution;
 KW fat distribution; male pattern; boar taint; flavour; impairment;
 KW reliable; immunocastration; meat production.

XX OS Chimeric - Mammalia.

OS Chimeric - Pasteurella haemolytica.

PN WO9956771-A2.

PD 11-NOV-1999.

PF 05-MAY-1999; 99WO-CA00360.

PR 05-MAY-1998; 98US-0084217.

PA (BIOS-) BIOSTAR INC.

PI Manns JG, Acres SD, Harland R;

XX WPI; 2000-062125/05.

DR N-PSDB; AAZ46400.

PT Production of uncastrated male food animals using vaccines -

XX Claim 22; Fig 3A-3F; 87pp; English.

XX This sequence represents a chimeric gonadotropin releasing
 CC hormone (GnRH)-leukotoxin (LKT) fusion protein, which may be
 CC used as a vaccine. The LKT portion of the protein acts to enhance
 CC the immunogenicity of the multimeric GnRH portion (AAY58135). The
 CC invention relates to a method of using two GnRH immunogen vaccines to
 CC produce uncastrated male animals for meat production, one vaccination
 CC prior to or during the fattening period to reduce circulating
 CC testosterone levels, and the second vaccination about 2-8 weeks before
 CC slaughter to substantially reduce androgenic and/or non-androgenic
 CC steroids. The invention is used to produce food animals that exhibit the
 CC weight gain and muscle/fat distribution of male animals without the
 CC problems associated with male animals. Such problems include "boar
 CC taint", a urine-like odour found in cooked meat of uncastrated pigs which
 CC is caused by steroids stored in the tissues, and similar flavour
 CC impairments in the meat of other intact male animals. The invention is

CC more reliable than prior art immunocastration techniques.

SQ Sequence 695 AA;
Query Match 1.1%; Score 10; DB 21; Length 695;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 359 LAQRVAAGLS 368
|||||

RESULT 20

ABB67818
ID ABB67818 standard; Protein; 802 AA.

XX AC ABB67818;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 30246.

XX KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL11921.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

XX PS Disclosure; SEQ ID NO 30246; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

XX CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 802 AA;

Query Match 1.1%; Score 10; DB 22; Length 802;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 353 AGVSAAVGS 362

Db 484 AGVSAAVGS 493
|||||

RESULT 21

AAR10889

ID AAR10889 standard; Protein; 924 AA.

XX AC AAR10889;

XX DT 11-APR-1991 (first entry)

XX DE Leukotoxin 352 encoded by plasmid PAA352.

XX KW LKT; vaccine; antigen; respiratory disease; shipping fever pneumonia.

XX OS Pasteurella haemolytica A1 strain B122.

XX PN CA2014033-A.

XX PD 07-OCT-1990.

XX PF 06-APR-1990; 90CA-2014033.

XX PR 07-APR-1989; 89US-0335018.

XX PA (UYSA-) UNIV SASKATCHEWAN.

XX PI Acres SD, Babiuk LA, Potter AA, Lawman MJP;

XX WPI; 1991-000097/01.

XX N-PSDB; AAQ10272.

XX PT Pasteurella haemolytica proteins and genes - used for producing vaccines to protect animals esp. cattle from respiratory diseases e.g. pneumonia.

XX PS Claim 13; Fig 5; 87pp; English.

XX CC Plasmid PAA352 is derived from PAA114, a clone isolated from a genomic library of P. haemolytica. The protein, designated "new leukotoxin" or "LKT 352" is 98% homologous to authentic leukotoxin. LKT 352 and pref. antigenic fragments of it, can be used in vaccines to protect cattle from respiratory diseases. They can also be used to produce antibodies for immunoaffinity purificn. of further proteins. [Fig. contg. sequence v. poor].
XX CC See also AAR10890, AAR10909, AAR10910 and AAQ10783.

XX SQ Sequence 924 AA;

Query Match 1.1%; Score 10; DB 12; Length 924;

Best Local Similarity 100.0%; Pred. No. 3.7;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274

Db 254 LAQRVAAGLS 263
|||||

RESULT 22

AAR42385

ID AAR42385 standard; Protein; 924 AA.

XX AC AAR42385;

XX DT 25-MAR-2003 (updated)

XX DT 19-APR-1994 (first entry)

XX DE Recombinant leukotoxin peptide from plasmid pCRR28.

XX KW Haemophilus somnus; immunogenic; haemolysin; LppB; LppC; thromboembolic meningencephalitis; septicaemia; arthritis; pneumonia; lktA gene; haemin-binding protein; fusion protein.

XX OS Pasteurella haemolytica.

XX PN WO9321323-A1.

```

XX 28-OCT-1993.
XX 05-APR-1993; 93WO-CA00135.
XX 09-APR-1992; 92US-0865050.
XX 04-JUN-1992; 92US-0893424.
XX 04-JUN-1992; 92US-0893426.
XX 29-MAR-1993; 93US-0038287.
XX 29-MAR-1993; 93US-0038288.
XX 29-MAR-1993; 93US-0038719.
XX (UYSA-) UNIV SASKATCHEWAN.
XX Harland RJ, Pfeiffer CG, Pontarollo RA, Potter AA;
XX Rioux C, Theisen M;
XX WPI; 1993-351733/44.
XX N-PSDB; AAQ51086.
XX Haemophilus somnus immunogenic proteins used in vaccines -
XX selected from haemin-binding protein, haemolysin, LppB and LppC,
XX and corresp. DNA
XX Disclosure; Fig 11; 119pp; English.
XX The lppB gene protein was expressed in E. coli as a fusion to the
XX Pasteurella haemolytica leukotoxin gene lktA coded for by plasmid
XX pAA352. The lppB gene fragment was taken from pMS11. LppB can be
XX used in vaccines for preventing or treating H. somnus infections,
XX which cause thromboembolic meningo-encephalitis, septicemia, arthritis
XX and pneumonia in vertebrates.
XX See also AAR42370-86.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 924 AA;
XX Query Match 1.1%; Score 10; DB 14; Length 924;
XX Best Local Similarity 100.0%; Pred. No. 3.7;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 265 LAQRVAAGLS 274
DB 256 LAQRVAAGLS 265
|||||

RESULT 23
AAR42380
ID AAR42380 standard; Protein; 924 AA.
XX AC AAR42380;
XX 25-MAR-2003 (updated)
DT 19-APR-1994 (first entry)
XX DE Recombinant leukotoxin peptide (split) from plasmid pGCH4.
XX KW Haemophilus somnus; immunogenic; haemolysin; LppB; LppC;
XX thromboembolic meningoencephalitis; septicemia; arthritis;
XX pneumonia; lktA gene; haemin-binding protein; fusion protein.
XX OS Pasteurella haemolytica.
XX PN WO9321323-A1.
XX PD 28-OCT-1993.
XX PF 05-APR-1993; 93WO-CA00135.
XX PR 09-APR-1992; 92US-0865050.
XX 04-JUN-1992; 92US-0893424.
XX 04-JUN-1992; 92US-0893426.
XX 29-MAR-1993; 93US-0038287.
XX 29-MAR-1993; 93US-0038288.
XX 29-MAR-1993; 93US-0038719.
XX (UYSA-) UNIV SASKATCHEWAN.
XX Harland RJ, Pfeiffer CG, Pontarollo RA, Potter AA;
XX Rioux C, Theisen M;

PR 29-MAR-1993; 93US-0038288.
PR 29-MAR-1993; 93US-0038719.
XX (UYSA-) UNIV SASKATCHEWAN.
XX Harland RJ, Pfeiffer CG, Pontarollo RA, Potter AA;
XX Rioux C, Theisen M;
XX WPI; 1993-351733/44.
XX N-PSDB; AAQ51082.
XX Haemophilus somnus immunogenic proteins used in vaccines -
XX selected from haemin-binding protein, haemolysin, LppB and LppC,
XX and corresp. DNA
XX Disclosure; Fig 6; 119pp; English.
XX The hmb gene encoding the haemin-binding protein was expressed in
XX E. coli as a fusion to the Pasteurella haemolytica leukotoxin gene
XX lktA coded for by plasmid pAA352. The hmb gene fragment was taken
XX from pRAP504 and starts at the codon for the 33 rd amino acid residue
XX of ORF1. The haemin binding protein can be used in vaccines for
XX preventing or treating H. somnus infections, which cause thromboembolic
XX meningo-encephalitis, septicemia, arthritis and pneumonia in
XX vertebrates.
XX See also AAR42370-86.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 924 AA;
XX Query Match 1.1%; Score 10; DB 14; Length 924;
XX Best Local Similarity 100.0%; Pred. No. 3.7;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 265 LAQRVAAGLS 274
DB 256 LAQRVAAGLS 265
|||||

RESULT 24
AAR42378
ID AAR42378 standard; Protein; 924 AA.
XX AC AAR42378;
XX 25-MAR-2003 (updated)
DT 19-APR-1994 (first entry)
XX DE Recombinant leukotoxin peptide (split) from plasmid pGCH5.
XX KW Haemophilus somnus; immunogenic; haemolysin; LppB; LppC;
XX thromboembolic meningoencephalitis; septicemia; arthritis;
XX pneumonia; lktA gene; haemin-binding protein; fusion protein.
XX OS Pasteurella haemolytica.
XX PN WO9321323-A1.
XX PD 28-OCT-1993.
XX PF 05-APR-1993; 93WO-CA00135.
XX PR 09-APR-1992; 92US-0865050.
XX 04-JUN-1992; 92US-0893424.
XX 04-JUN-1992; 92US-0893426.
XX 29-MAR-1993; 93US-0038287.
XX 29-MAR-1993; 93US-0038288.
XX 29-MAR-1993; 93US-0038719.
XX (UYSA-) UNIV SASKATCHEWAN.
XX Harland RJ, Pfeiffer CG, Pontarollo RA, Potter AA;
XX Rioux C, Theisen M;

```

XX WPI; 1993-351733/44.
 DR N-PSDB; AAQ51081.
 XX Haemophilus somnus immunogenic proteins used in vaccines -
 PT selected from haemin-binding protein, haemolysin, LppB and LppC,
 PT and corresp. DNA
 XX Disclosure; Fig 5; 119pp; English.
 XX The hmb gene encoding the haemin-binding protein was expressed in
 CC E. coli as a fusion to the Pasteurella haemolytica leukotoxin gene
 CC lktA coded for by plasmid pAA352. The hmb gene fragment was taken
 CC from pRAP501 and starts at the codon for the third amino acid residue
 CC of ORF1. The haemin binding protein can be used in vaccines for
 CC preventing or treating H. somnus infections, which cause thromboembolic
 CC meningo-encephalitis, septicemia, arthritis and pneumonia in
 CC vertebrates.
 CC See also AAR42370-86.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 924 AA;
 SQ

Query Match 1.1%; Score 10; DB 14; Length 924;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQRVAAGLS 274
 DB 256 LAQRVAAGLS 265
 |||||
 |||||

RESULT 25
 AAR1482
 ID AAR14482 standard; Protein; 926 AA.
 AC AAR14482;
 XX 15-JAN-1992 (first entry)
 DT LKT352.
 DE
 XX Antigen; leukotoxin; vaccine; lktA.
 XX Pasteurella haemolytica.
 XX WO9115237-A.
 XX 17-OCT-1991.
 XX 17-OCT-1991; 91WO-CA00170.
 XX 05-APR-1990; 90US-0504850.
 XX (UYSA-) UNIV SASKATCHEWAN.
 XX Acres SD, Bariuk LA, Potter AA, Lawman MJP;
 XX WPI; 1991-324967/44.
 XX Vaccines for Pasteurella haemolytica infection in cattle -
 PT comprise sub-unit antigens from P haemolytica fibrillar protein,
 PT plasmin receptor, 50 K outer membrane protein and leukotoxin.
 XX Disclosure; Fig 5; 92pp; English.
 XX LKT352 is 98% homologous with authentic leukotoxin and migrates
 CC to the same position on gels.
 CC The LKT352 gene was prep'd. as follows: lktA, an MaeI fragment
 CC contg. the gene was ligated into the SmaI site of pUC13 to form
 CC pAA179. From this, two constructs were made in the pTAC-based
 CC vector, pGH432:lacI digested with SmaI. One, pAA342, consisted of
 CC the 5' AbaIII fragment from lktA while the other, pAA345, contained

CC the entire MaeI fragment. Clone pAA342 expressed a truncated
 CC leukotoxin peptide at high levels while pAA345 expressed full
 CC length leukotoxin at very low levels. The 3' end of the lktA gene
 CC of pAA345 was therefore ligated to StyI/BamHI digested pAA342 to
 CC yield pAA352 contg. the LKT352 sequence. The protein expressed
 CC from the vector can be used to prepare a subunit vaccine with
 CC other P. haemolytica antigens, e.g. fimbrial protein, plasmin
 CC receptor or 50K outer membrane protein. The vaccines can be used
 CC to protect cattle from respiratory diseases such as pneumonia, esp.
 CC shipping fever pneumonia.
 CC See also AAR14481, 83,84 and 85.
 XX
 XX Sequence 926 AA;
 SQ

Query Match 1.1%; Score 10; DB 12; Length 926;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQRVAAGLS 274
 DB 256 LAQRVAAGLS 265
 |||||
 |||||

RESULT 26
 AAR34545
 ID AAR34545 standard; Protein; 926 AA.
 XX
 AC AAR34545;
 XX 25-MAR-2003 (updated)
 DT 23-AUG-1993 (first entry)
 XX Leukotoxin 352 produced from pAA352.
 DE
 XX Vector; LKT 352; flanking; recombinant; antigen; somatostatin;
 KW gonadotropin releasing hormone; rotavirus viral protein 4;
 KW carrier protein; lactation; reproduction.
 XX
 OS Pasteurella haemolytica.
 XX WO9308290-A1.
 XX 29-APR-1993.
 XX 15-OCT-1992; 92WO-CA00449.
 XX 16-OCT-1991; 91US-0779171.
 XX 14-OCT-1992; 92US-0960932.
 XX (UYSA-) UNIV SASKATCHEWAN.
 XX Hughes HPA, Potter AA, Redmond MJ;
 XX WPI; 1993-152482/18.
 XX N-PSDB; AAQ41317.
 XX Immunological carrier system with enhanced immunogenicity -
 PT comprises chimeric protein comprising leuco:toxin peptide or
 PT homologous protein fused to antigen esp. somatostatin or
 PT gonadotropin releasing hormone
 XX Disclosure; Fig 3; 95pp; English.
 XX Gene libraries of P. haemolytica A1 (strain Bl22) were constructed
 CC in lambda gt11 and pUC13. Resulting clones were used to transform E.
 CC coli and individual colonies were pooled and screened for reaction
 CC with serum from a calf which had survived a P. haemolytica infection
 CC and that had been boosted with a conc. culture supernatant of P.
 CC haemolytica to increase anti-leukotoxin antibody levels. Positive
 CC colonies were screened for their ability to produce leukotoxin by
 CC incubating cell lysates with bovine neutrophils and measuring the
 CC release of lactate dehydrogenase from the neutrophils. A 4kb
 CC fragment was obtd. Progressively larger clones were isolated by

CC chromosome walking to isolate full length recombinants of ca. 8kb.
 CC in pAA114. The clone was subjected to restriction enzyme digestion
 CC to yield two clones, one expressing truncated leukotoxin peptide at
 CC high levels and the other expressing the full length leukotoxin at
 CC low levels. The 3' end of the lktA gene from the full length clone
 CC was ligated to the truncated gene clone to yield plasmid pAA352. The
 CC clone was used to produce chimeric proteins by gene fusion with an
 CC antigen coding sequence, e.g. the coding sequence of somatostatin,
 CC gonadotropin releasing hormone or rotavirus viral protein 4, i.e.
 CC leukotoxin works as a carrier protein to bring about a larger
 CC immune response than the antigen alone. Immunisation with these
 CC antigens can regulate growth rate, lactation and reproductive
 CC efficiency. See also AAR34546-8.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 926 AA;
 Query Match 1.1%; Score 10; DB 14; Length 926;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 265 LAQ RVAAGLS 274
 DB 256 LAQ RVAAGLS 265
 |||||
 RESULT 27
 AAR50291
 ID AAR50291 standard; Protein; 926 AA.
 XX
 AC AAR50291;
 XX
 DT 06-OCT-1994 (first entry)
 XX
 DE Recombinant leukotoxin from plasmid pAA352.
 XX
 KW Vaccine; outer membrane protein; OMP; Haemophilus somnus;
 KW iron regulated protein; leukotoxin; Pasteurella haemolytica;
 KW LKT352.
 XX
 OS Pasteurella haemolytica A1 (strain B122).
 XX
 PN CA2099707-A.
 PD 03-JAN-1994.
 XX
 PF 29-JUN-1993; 93CA-2099707.
 XX
 PR 02-JUL-1992; 92US-0908253.
 XX
 PA (UYSA-) UNIV SASKATCHEWAN.
 XX
 PI Harland RJ, Potter AA;
 DR WPI; 1994-092909/12.
 DR N-PSDB; AAQ44760.
 XX
 PT Haemophilus somnus outer membrane protein extract -
 PT enriched with iron-regulated proteins, opt. contg.
 PT leuco:toxin antigens, for use as vaccine
 XX
 PS Claim 5; Fig 5; 78pp; English.
 XX
 CC A vaccine comprising an outer membrane protein (OMP) extract of
 CC Haemophilus somnus enriched with iron regulated proteins is new.
 CC The vaccine pref. further comprises an immunogenic leukotoxin
 CC polypeptide, esp. an immunogenic Pasteurella haemolytica leukotoxin
 CC homologous to LKT352. Example 1.2 describes the prodn. of
 CC P. haemolytica recombinant leukotoxin from pAA352.
 CC Two expression constructs were made. One, pAA342, contained the
 CC 5'-Ahali fragment of the lktA gene, while the other, pAA345,
 CC contained the entire lktA gene. pAA342 expressed a truncated
 CC leukotoxin peptide at high levels, while pAA345 expressed full

CC length leukotoxin at very low levels. Therefore, the 3' end
 CC of the lktA gene was ligated into pAA342, yielding plasmid pAA352.
 CC LKT352 or new leukotoxin is 98% homologous to authentic
 CC leukotoxin.
 CC NB: the protein sequence in Fig 5 comprises 926 amino acids,
 CC however this protein is described in the text as having
 CC 931 amino acids.
 XX
 SQ Sequence 926 AA;
 Query Match 1.1%; Score 10; DB 15; Length 926;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 265 LAQ RVAAGLS 274
 DB 256 LAQ RVAAGLS 265
 |||||
 RESULT 28
 AAW03945
 ID AAW03945 standard; Protein; 926 AA.
 XX
 AC AAW03945;
 XX
 DT 20-NOV-1996 (first entry)
 XX
 DE P. haemolytica truncated leukotoxin (LKT352).
 XX
 KW Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
 KW fusion protein; immunogen; vaccine; fertility control;
 KW contraceptive; sterilisation; plasmid pAA352.
 XX
 OS Pasteurella haemolytica A1 strain B122.
 XX
 PN WO9624675-A1.
 PD 15-AUG-1996.
 XX
 PF 24-JAN-1996; 96WO-CA00049.
 XX
 PR 10-FEB-1995; 95US-0387156.
 XX
 PA (UYSA-) UNIV SASKATCHEWAN.
 XX
 PI Manns JG, Potter AA;
 DR WPI; 1996-384447/38.
 DR N-PSDB; AAT37179.
 XX
 PT Gonadotropin-releasing hormone multimer fusion proteins - with
 PT leukotoxin polypeptide for increased immunogenicity, useful in
 PT antifertility vaccine prodn.
 XX
 PS Example 1; Fig 3A-3I; 87pp; English.
 XX
 CC A truncated leukotoxin (AAW03945), LKT 352, lacks the cytotoxic
 CC portion of the native protein from Pasteurella haemolytica. It is
 CC the product of plasmid pAA352 which carries a truncated lktA gene
 CC (AAT37179). A fusion protein (AAW03942) between LKT352 and a
 CC gonadotropin releasing hormone tetramer can be expressed in
 CC Escherichia coli. This is useful as a vaccine for fertility
 CC control, partic. immunological sterilisation of domestic or
 CC farm animals.
 XX
 SQ Sequence 926 AA;
 Query Match 1.1%; Score 10; DB 17; Length 926;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 265 LAQ RVAAGLS 274
 |||||

```

Db          256 LAQRVAAGLS 265

RESULT 29
AAW79568
ID AAW79568 standard; Protein; 926 AA.
XX
XX
AC AAW79568;
XX
XX
DT 24-DEC-1998 (first entry)
XX
XX Leukotoxin 352 polypeptide.
XX
XX Gonadotropin releasing hormone; GnRH; chimera; leukotoxin polypeptide;
KW multimer; vaccine; tumour; Leukotoxin 352; LKT 352; lKtA; plasmid pAA352;
KW cytotoxic.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Key 11..924
FT Misc_feature /note= "Recombinant leukotoxin peptide"
XX
XX WO9806848-A1.
XX
XX 19-FEB-1998.
XX
XX 08-AUG-1997; 97WO-CA00559.
XX
XX 09-AUG-1996; 96US-0694865.
XX
XX (UUSA-) UNIV SASKATCHEWAN.
XX
XX Manns JG, Potter AA;
XX
XX WPI; 1998-159540/14.
DR N-PSDB; AAW61530.
XX
XX Chimeric protein of leukotoxin and gonadotropin releasing hormone -
PT useful for, e.g. preparation of vaccines for reduction of incidence
PT of mammary tumours in mammals
XX
XX Claim 7; Figure 3.1-9; 118pp; English.
XX
XX The present sequence represents a recombinantly produced or chemically
CC synthesised leukotoxin 352 (LKT 352)polypeptide, derived from the lKtA
CC gene that is present in the plasmid pAA352. This gene produces a
CC truncated protein that has an estimated molecular weight of about 99 kDa
CC and lacks the cytotoxic portion of the molecule. Thus this gene has a
CC higher expression level than that of the full-length molecule. This can
CC be used in the construction of a chimeric protein that comprises a
CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The
CC chimeric protein can be used as a vaccine to help reduce the incidence of
CC mammary tumours in a mammalian individual.
XX
XX Query Match 1.1%; Score 10; DB 19; Length 926;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 265 LAQRVAAGLS 274
|||||||
Db 256 LAQRVAAGLS 265

RESULT 30
AAW07637
ID AAW07637 standard; Protein; 934 AA.
XX
XX AAW07637;
XX
XX 25-MAR-2003 (updated)

```

```

DT XX 05-FEB-1997 (first entry)
DE XX
XX P. suis leukotoxin gene pslktA, protein product.
KW Leukotoxin; pslktC; pslktA; pslktB; calcium dependent; RTX;
KW cytotoxic protein; repeat of toxin; BL-3 cell; pig; lymphocyte;
KW immunogen; vaccine; immunisation; diarrhoea; reagent;
KW diagnosis; assay; enteritis.
XX
OS Pasteurella suis.
XX
XX Key Location/Qualifiers
FH Key 229..245
FT Domain /label= transmembrane_domain
FT Domain 235..311
FT Domain /label= transmembrane_domain
FT Domain 377..393
FT Domain /label= transmembrane_domain
FT Region 710..718
FT Region /note= "glycine rich repeat sequence"
FT Region 719..727
FT Region /note= "glycine rich repeat sequence"
FT Region 728..736
FT Region /note= "glycine rich repeat sequence"
FT Region 737..745
FT Region /note= "glycine rich repeat sequence"
FT Region 746..754
FT Region /note= "glycine rich repeat sequence"
FT Region 756..763
FT Region /note= "glycine rich repeat sequence"
FT Region 764..772
FT Region /note= "glycine rich repeat sequence"
FT Region 773..781
FT Region /note= "glycine rich repeat sequence"
XX
XX US5559008-A.
PN XX
XX
PD 24-SEP-1996.
XX
XX 22-MAR-1994; 94US-0215805.
PF
XX
XX 05-NOV-1992; 92US-0972157.
PR
XX 22-MAR-1994; 94US-0215805.
XX
XX (CORR ) CORNELL RES FOUND INC.
PA
XX Chang Y;
XX
XX WPI; 1996-442408/44.
DR N-PSDB; AAT45417.
XX
XX Leukotoxin genes from Pasteurella suis - also vectors and cells for
PT expressing gene prods., for use in vaccines against porcine
PT enteritis
XX
XX Claim 34; Columns 59-64; 47pp; English.
XX
XX The present sequence is protein prod. of the Pasteurella suis
CC leukotoxin gene pslktA, which together with the pslktC, and pslktB
CC gene prods., comprises a high mol. wt. calcium dependent cytotoxic
CC protein of the repeat of toxin family (RTX), capable of killing
CC both BL-3 cells and pig lymphocytes. The leukotoxin proteins, and
CC the pslktC, A and B genes, or fragments, can be used as immunogens
CC in vaccines for the immunisation of pigs against diarrhoea, or as
CC reagents in diagnostic assays. The leukotoxin genes were isolated
CC by screening a P. suis genomic library, in the phage vector
CC lambda-dash, with a DNA probe derived from pYFC19, a plasmid
CC carrying the lKtCA locus (Chang et al., Infect. Immun., vol. 55,
CC pp2348-2354, 1987).
XX
XX (Updated on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 934 AA;
SQ

```

Query Match 1.1%; Score 10; DB 17; Length 934;
 Best Local Similarity 100.0%; Pred. No. 3.8; 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
 |||||
 Db 279 LAQRVAAGLS 288

RESULT 31
 AAR34547
 ID AAR34547 standard; Protein; 936 AA.

XX AC AAR34547;
 XX DT 25-MAR-2003 (updated)
 XX DT 23-AUG-1993 (first entry)
 XX GNRH-leukotoxin gene fusion prod.
 XX Vector; LKT 352; flanking; recombinant; antigen; somatostatin;
 KW gonadotropin releasing hormone; rotavirus viral protein 4;
 KW carrier protein; lactation; reproduction; SRIF.
 XX OS Synthetic.

XX FH Key Location/Qualifiers
 XX FT Protein 1..926
 FT /note= "recombinant leukotoxin protein"
 FT Peptide 927..936
 FT /note= "GNRH"

XX PN WO9308290-A1.
 XX PD 29-APR-1993.
 XX PF 15-OCT-1992; 92WO-CA00449.
 XX PR 16-OCT-1991; 91US-0779171.
 XX PR 14-OCT-1992; 92US-0960932.
 XX PA (UYSA-) UNIV SASKATCHEWAN.

XX PI Hughes HPA, Potter AA, Redmond MJ;
 XX DR WPI; 1993-152482/18.
 DR N-PSDB; AAQ41322.

XX Immunological carrier system with enhanced immunogenicity -
 PT comprises chimeric protein comprising leuco:toxin peptide or
 PT homologous protein fused to antigen esp. somatostatin or
 PT gonadotropin releasing hormone

XX Example 2; Fig 8; 95pp; English.

XX Oligonucleotides contg. sequences from bovine gonadotropin
 CC releasing hormone (GNRH) gene were constructed on a Pharmacia Gene
 CC Assembler using standard phosphoramidite chemistry. The oligo-
 CC nucleotides were annealed and ligated into vector pAA352 (contg.
 CC the Pateurella heamolytica leuco-toxin gene) which had been digested
 CC with BamHI. The ligated DNA was used to transform E. coli strain
 CC MH3000. Transformants contg. the oligonucleotide inserts were
 CC identified by restriction endonuclease mapping and the recombinant
 CC plasmid designated PAA502. The chimeric protein produced from the
 CC plasmid works to bring about a larger immune response than the antigen
 CC alone, i.e. the leukotoxin works as a carrier protein.
 CC See also AAR34545-8.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 936 AA;

Query Match 1.1%; Score 10; DB 14; Length 936;
 Best Local Similarity 100.0%; Pred. No. 3.8;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 265 LAQRVAAGLS 274
 |||||
 Db 256 LAQRVAAGLS 265

RESULT 32

AAR34546
 ID AAR34546 standard; Protein; 943 AA.

XX AC AAR34546;
 XX DT 25-MAR-2003 (updated)
 XX DT 23-AUG-1993 (first entry)
 XX Somatostatin-leukotoxin gene fusion prod.
 XX Vector; LKT 352; flanking; recombinant; antigen; somatostatin;
 KW gonadotropin releasing hormone; rotavirus viral protein 4;
 KW carrier protein; lactation; reproduction; SRIF.
 XX OS Synthetic.

XX FH Key Location/Qualifiers
 XX FT Protein 1..929
 FT /note= "recombinant leukotoxin protein"
 FT Peptide 930..943
 FT /note= "SRIF"

XX PN WO9308290-A1.
 XX PD 29-APR-1993.
 XX PF 15-OCT-1992; 92WO-CA00449.
 XX PR 16-OCT-1991; 91US-0779171.
 XX PR 14-OCT-1992; 92US-0960932.
 XX PA (UYSA-) UNIV SASKATCHEWAN.

XX PI Hughes HPA, Potter AA, Redmond MJ;
 XX DR WPI; 1993-152482/18.
 DR N-PSDB; AAQ41321.

XX Immunological carrier system with enhanced immunogenicity -
 PT comprises chimeric protein comprising leuco:toxin peptide or
 PT homologous protein fused to antigen esp. somatostatin or
 PT gonadotropin releasing hormone

XX Example 2; Fig 6; 95pp; English.

XX Oligonucleotides contg. sequences from bovine somatostatin (SRIF)
 CC gene were constructed on a Pharmacia Gene Assembler using standard
 CC phosphoramidite chemistry. The oligonucleotides were annealed and
 CC ligated into vector pAA352 (contg. the Pateurella heamolytica leuco-
 CC toxin gene) which had been digested with BamHI. The ligated DNA was
 CC used to transform E. coli strain JM105. Transformants contg. the
 CC oligonucleotide inserts were identified by restriction endonuclease
 CC mapping and the recombinant plasmid designated pAA496. The
 CC chimeric protein produced from the plasmid works to bring about a
 CC larger immune response than the antigen alone, i.e. the leukotoxin
 CC works as a carrier protein.
 CC See also AAR34545-8.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 943 AA;

Query Match 1.1%; Score 10; DB 14; Length 943;
 Best Local Similarity 100.0%; Pred. No. 3.8;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQ RVAAGLS 274
 |||||
 Db 256 LAQ RVAAGLS 265

RESULT 33
 AAR34548
 ID AAR34548 standard; Protein; 951 AA.

XX AAR34548;
 AC AAR34548;
 XX 25-MAR-2003 (updated)
 DT 23-AUG-1993 (first entry)

XX Rotavirus VP4-leukotoxin gene fusion prod.

XX Vector; LKT 352; flanking; recombinant; antigen; somatostatin;
 KW gonadotropin releasing hormone; rotavirus viral protein 4;
 KW carrier protein; lactation; reproduction; SRIF.

XX Synthetic.

XX Key Location/Qualifiers
 FH Protein 1..926
 FT /note= "recombinant leukotoxin protein"
 FT Peptide 927..951
 FT /note= "Rotavirus VP-4"

XX WO9308290-A1.

PN 29-APR-1993.

XX 15-OCT-1992; 92WO-CA00449.

XX 16-OCT-1991; 91US-0779171.

XX 14-OCT-1992; 92US-0960932.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Hughes HPA, Potter AA, Redmond MJ;

XX WPI; 1993-152482/18.

DR N-PSDB; AAQ41322.

XX Immunological carrier system with enhanced immunogenicity -
 PT comprises chimeric protein comprising leuco:toxin peptide or
 PT homologous protein fused to antigen esp. somatostatin or
 PT gonadotropin releasing hormone

XX Example 2; Fig 8; 95pp; English.

XX Oligonucleotides contg. sequences from bovine Rotavirus viral
 CC protein 4 (VP-4) gene were constructed on a Pharmacia Gene
 CC Assembler using standard phosphoramidite chemistry. The oligo-
 CC nucleotides were annealed and ligated into vector pAA352 (contg.
 CC the Pasteurella haemolytica leuko-toxin gene) which had been digested
 CC with BamHI. The ligated DNA was used to transform E. coli strain
 CC MH3000. Transformants contg. the oligonucleotide inserts were
 CC identified by restriction endonuclease mapping and the recombinant
 CC plasmid designated pAA501. The chimeric protein produced from the
 CC plasmid works to bring about a larger immune response than the antigen
 CC alone, i.e. the leukotoxin works as a carrier protein.

CC See also AAR34545-7;
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 951 AA;

Query Match 1..18; Score 10; DB 14; Length 951;
 Best Local Similarity 100.0%; Pred. No. 3.8;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQ RVAAGLS 274
 |||||

Db 256 LAQ RVAAGLS 265

RESULT 34

AAR07167

ID AAR07167 standard; protein; 953 AA.

XX AAR07167;

XX 24-JAN-1991 (first entry)

XX 105kd PTX protein of P.haemolytica.

XX Vaccine; shipping fever.

XX Pasteurella haemolytica.

XX US4957739-A.

XX 18-SEP-1990.

XX 13-AUG-1987; 87US-0085430.

XX 13-AUG-1987; 87US-0085430.

XX (TEXA) UNIV OF TEXAS SYSTE.

XX Berget P, Engler M, Highlander S, Weinstock G;

XX WPI; 1990-304558/40.

DR N-PSDB; AAQ06074.

XX Vaccine against shipping fever in cattle - contains purified
 PT Pasteurella haemolytica antigen of molecular wt. 105 kD

XX Disclosure; Fig 9a-h; 35pp; English.

XX Gene product is antigenic to P.haemolytica, and may be used as a
 CC vaccine for immunisation against shipping fever. Abs raised to the
 CC antigen may be used in passive immunisation and diagnosis.

XX SQ Sequence 953 AA;

Query Match 1..18; Score 10; DB 11; Length 953;
 Best Local Similarity 100.0%; Pred. No. 3.8;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQ RVAAGLS 274
 |||||

Db 283 LAQ RVAAGLS 292

RESULT 35

AAR15159

ID AAR15159 standard; Protein; 953 AA.

XX AAR15159;

XX 10-JAN-1992 (first entry)

XX Leukotoxin from P. haemolytica.

XX Leukotoxin; probe; pasteurellosis; cytotoxicity; leukocytes.

XX Pasteurella haemolytica A1 (ATCC 43270).

XX US5055400-A.

XX 08-OCT-1991.

XX 26-NOV-1986; 86US-0935493.

XX 26-NOV-1986; 86US-0935493.


```

RESULT 38
AAE04638
ID AAE04638 standard; Protein; 953 AA.
XX
AC AAE04638;
XX
AC AAE04638;
XX
DT 10-SEP-2001 (first entry)
XX
DE Pasteurella haemolytica leukotoxin protein.
XX
KW Leukotoxin; respiratory disease; infection; therapy; immunostimulant;
KW antibacterial; vaccine; transgenic plant; animal feed.
XX
OS Pasteurella haemolytica.
XX
FH Key Location/Qualifiers
FT Domain 230..250
FT /label= Hydrophobic_transmembrane_domain_1
FT Domain 280..320
FT /label= Hydrophobic_transmembrane_domain_2
FT Domain 360..400
FT /label= Hydrophobic_transmembrane_domain_3
XX
PN WO200144289-A2.
XX
PD 21-JUN-2001.
XX
PF 15-DEC-2000; 2000WO-CA01498.
XX
PR 17-DEC-1999; 99US-0172148.
XX
PA (UYGU-) UNIV GUELPH OFFICE.
XX
PI Lo RVC, Shewen PE, Lee RWH, Hodgins D, Strommer JN;
XX WPI; 2001-408470/43.
XX
XX Modified leukotoxin polypeptide is useful in a vaccine to prevent or
XX treat Mannheimia (Pasteurella) infection (particularly M. haemolytica
XX infection), and disease associated with a leukotoxin, e.g., respiratory
XX disease -
XX
PS Claim 6; Fig 13; 70pp; English.
XX
XX The present sequence is Pasteurella (Mannheimia) haemolytica full
XX length leukotoxin protein. The present invention relates to modified
XX leukotoxin DNA sequences, wherein the modification comprises the
XX removal of DNA sequence encoding the amino acids within the hydrophobic
XX transmembrane domain of a full length leukotoxin protein. Modified
XX leukotoxin sequences are used in vaccines to treat or prevent diseases
XX associated with leukotoxin, e.g., respiratory disease, and Mannheimia
XX infection (particularly M. haemolytica infection). In addition, the
XX vaccine is used to prepare a medicament. Furthermore, the plant
XX transformed with modified leukotoxin sequences is fed to an animal
XX such as a ruminant, to prevent or treat respiratory diseases.
XX
SQ Sequence 953 AA;
XX
Query Match 1.1%; Score 10; DB 22; Length 953;
Best Local Similarity 100.0%; Pred. No. 3.8; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;
XX
QY 265 LAQRVAAGLS 274
DB 283 LAQRVAAGLS 292
XX
RESULT 39
AAW03942
ID AAW03942 standard; Protein; 977 AA.
XX

```

```

AC AAW03942;
XX
DT 20-NOV-1996 (first entry)
XX
DE LKT-GnRH protein fusion from pCB113.
XX
KW Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
KW fusion protein; immunogen; vaccine; fertility control;
KW contraceptive; sterilisation.
XX
OS Chimeric Pasteurella haemolytica A1 strain B122;
OS Chimeric synthetic.
XX
FH Key Location/Qualifiers
FT Domain 1..929
FT /label= LKT
FT Domain 927..977
FT /label= GnRH_repeat_domain
XX
PN WO9624675-A1.
XX
PD 15-AUG-1996.
XX
PF 24-JAN-1996; 96WO-CA00049.
XX
PR 10-FEB-1995; 95US-0387156.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX
PI Manns JG, Potter AA;
XX WPI; 1996-384447/38.
XX N-PSDB; AAT37176.
XX
XX Gonadotropin-releasing hormone multimer fusion proteins - with
XX leukotoxin polypeptide for increased immunogenicity, useful in
XX antifertility vaccine prodn.
XX
PS Claim 7; Fig 5A-5H; 87pp; English.
XX
XX A chimeric protein (AAW03942) is composed of a fusion between
XX a truncated leukotoxin (LKT-352) from Pasteurella haemolytica (see
XX also AAW03945) and a 4-copy gonadoliberin-releasing hormone (GnRH)
XX repeat sequence (see also AAW03944). It is the product of a
XX chimeric gene (AAT37176) produced by ligating a synthetic sequence
XX for the 4-copy GnRH into vector pAA352 (ATCC 68283), which carries
XX the LKT-352 gene. Recombinant plasmid pCB113 (LKT 352:4 copy
XX GnRH, ATCC 69749) was obt'd. Escherichia coli transformants
XX produced the chimeric protein, which is useful as a vaccine for
XX fertility control, esp. immunological sterilisation of
XX domestic or farm animals.
XX
SQ Sequence 977 AA;
XX
Query Match 1.1%; Score 10; DB 17; Length 977;
Best Local Similarity 100.0%; Pred. No. 3.9; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;
XX
QY 265 LAQRVAAGLS 274
DB 256 LAQRVAAGLS 265
XX
RESULT 40
AAW79569
ID AAW79569 standard; Protein; 977 AA.
XX
AC AAW79569;
XX
DT 24-DEC-1998 (first entry)
XX
DE LKT-GnRH chimeric protein.
XX

```

KW Chimera; pCB113; LKT 352; GnRH; Gonadotropin releasing hormone; multimer;
KW cytotoxic activity; antigen presentation; immune response; vaccine;
KW tumour.
XX Synthetic.
OS
XX WO9806848-A1.
XX
XX
XX
XX
PD 19-FEB-1998.
XX
XX 08-AUG-1997; 97WO-CA00559.
PF
XX
XX 09-AUG-1996; 96US-0694865.
PR
XX
XX (UYSA-) UNIV SASKATCHEWAN.
PA
XX
XX Manns JG, Potter AA;
PI
XX
XX WPI; 1998-159540/14.
DR
XX N-PSDB; AAV61531.
XX
XX
PT Chimeric protein of leukotoxin and gonadotropin releasing hormone -
PT useful for, e.g. preparation of vaccines for reduction of incidence
PT of mammary tumours in mammals
XX
PS Disclosure; Figure 5.1-8; 118pp; English.
XX
XX The present sequence represents the LKT-GnRH chimeric protein from
CC pCB113. This plasmid contains the LKT 352 polypeptide (AAW79568) fused
CC to four copies of the GnRH peptide. This chimera lacks cytotoxic
CC activity which enables there to be an increase in antigen presentation
CC and thus an optimal immune response. The removal of this region also
CC enables the truncated LKT to be expressed at much higher levels and
CC allows the amount of antigen administered to be reduced. This chimeric
CC protein comprises a leukotoxin polypeptide, several multimers, and a GnRH
CC sequence. The chimeric protein can be used as a vaccine to help reduce
CC the incidence of mammary tumours in a mammalian individual.
XX
SQ Sequence 977 AA;

Query Match 1.1%; Score 10; DB 19; Length 977;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAGLS 274
| | | | | | | |
Db 256 LAQRVAGLS 265

Search completed: February 17, 2004, 10:19:35
Job time : 66 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: February 17, 2004, 10:09:50 ; Search time 17 Seconds

(without alignments)
2564.339 Million cell updates/sec

Title: US-10-069-799-5

Perfect score: 4647

Sequence: 1 MSNINVKNIQAGLNSTKS.....SSNALQPIQTQTOGILAPSV 927

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|--------------|--------------------|
| 1 | 2334 | 50.2 | 953 | 1 LKAL_PASHA | P16535 pasteurella |
| 2 | 2332 | 50.2 | 953 | 1 LKAB_PASHA | P55118 pasteurella |
| 3 | 2306 | 49.6 | 947 | 1 LKTA_PASSP | P55123 pasteurella |
| 4 | 2256 | 48.5 | 953 | 1 LKAJ_PASHA | P55116 pasteurella |
| 5 | 2236.5 | 48.1 | 956 | 1 RT2A_ACTPL | P15377 actinobacil |
| 6 | 2236 | 48.1 | 955 | 1 LKAA_PASHA | P55117 pasteurella |
| 7 | 2214.5 | 47.7 | 956 | 1 HLYA_ACTSU | Q00951 actinobacil |
| 8 | 1955.5 | 42.1 | 1052 | 1 RT31_ACTPL | P55131 actinobacil |
| 9 | 1924 | 41.4 | 1049 | 1 HLYA_ECOLI | P08715 escherichia |
| 10 | 1848.5 | 39.8 | 1024 | 1 HLYA_ECOLI | P09983 escherichia |
| 11 | 1845 | 39.7 | 1023 | 1 HLY1_ECOLI | P55129 actinobacil |
| 12 | 1729.5 | 37.2 | 1023 | 1 RT12_ACTPL | P55128 actinobacil |
| 13 | 1723 | 37.1 | 1023 | 1 RT11_ACTPL | P16462 actinobacil |
| 14 | 1712.5 | 36.9 | 1050 | 1 LKTA_ACTAC | P15318 bordetella |
| 15 | 802 | 17.3 | 1705 | 1 CYAA_BORPE | Q57506 bordetella |
| 16 | 801.5 | 17.2 | 1705 | 1 CYAA_BORBR | Q57506 bordetella |
| 17 | 437 | 9.4 | 1829 | 1 FRPC_NEIMB | P55127 neisseria m |
| 18 | 435 | 9.4 | 1829 | 1 FRPC_NEIMC | Q9K0K9 neisseria m |
| 19 | 421 | 9.1 | 1302 | 1 FRPA_NEIMB | P55126 neisseria m |
| 20 | 408.5 | 8.8 | 1115 | 1 FRPA_NEIMC | P55126 neisseria m |
| 21 | 189.5 | 4.1 | 1608 | 1 HLYA_SERMA | P15320 serratia ma |
| 22 | 188.5 | 4.1 | 2869 | 1 RBPI_PLAVB | Q00798 plasmidium |
| 23 | 183.5 | 3.9 | 1577 | 1 HLYA_PROMI | P16466 proteus mir |
| 24 | 180.5 | 3.9 | 1683 | 1 YL24_ANASP | Q9YV57 anabaena sp |
| 25 | 178.5 | 3.8 | 284 | 1 NODO_RHILV | P15728 rickobium l |
| 26 | 178.5 | 3.8 | 1025 | 1 SLAP_CAUCR | P35828 caulobacter |
| 27 | 175.5 | 3.8 | 1228 | 1 SIAP_BACST | P35825 bacillus st |
| 28 | 173 | 3.7 | 1251 | 1 RBPI_PLAVB | Q00799 plasmidium |
| 29 | 170 | 3.7 | 491 | 1 ZAPA_PROMI | Q11137 proteus mir |
| 30 | 169.5 | 3.6 | 2249 | 1 ONPA_RICRI | P15921 rickettsia |
| 31 | 169 | 3.6 | 1520 | 1 PMPD_CHLMU | Q9PLB0 chlamydia m |
| 32 | 167 | 3.6 | 1656 | 1 OMPB_RICJA | O06653 r outer mem |
| 33 | 165 | 3.6 | 479 | 1 PRTC_ERWCH | P16317 erwinia chr |

RESULT 1

LKAL_PASHA
ID LKAL_PASHA STANDARD; PRT; 953 AA.
AC P16535;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Leukotoxin from serotype A1.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype A1;
RX MEDLINE=87306837; PubMed=3040588;
RA Lo R.Y.C., Strathdee C.A., Shewen P.E.;
RT "Nucleotide sequence of the leukotoxin genes of Pasteurella haemolytica A1";
RL Infect. Immun. 55:1987-1996(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype A1 / PHL101;
RX MEDLINE=89210283; PubMed=2707120;
RA Highlander S.K., Chidambaram M., Engler M.J., Weinstock G.M.;
RT "DNA sequence of the Pasteurella haemolytica leukotoxin gene cluster";
RL DNA 8:15-28(1989).
RN [3]
RP SEQUENCE OF 884-953 FROM N.A.
RC STRAIN=Serotype A1 / PHL101;
RX MEDLINE=90236888; PubMed=2185213;
RA Highlander S.K., Engler M.J., Weinstock G.M.;
RT "Secretion and expression of the Pasteurella haemolytica Leukotoxin";
RL J. Bacteriol. 172:2343-2350(1990).
CC -1- FUNCTION: Bacterial hemolysins are exotoxins that attack blood cell membranes and cause cell rupture by mechanisms not clearly defined.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DOMAIN: The Gly-rich region is probably involved in binding calcium, which is required for target cell-binding or cytolytic activity.
CC -1- DOMAIN: The three transmembrane domains are believed to be involved in pore formation by the cytotoxin (By similarity).
CC -1- PTM: Palmitoylated by lktC. The toxin only becomes active when modified (By similarity).
CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

Q07162 erwinia chr
P27470 streptococc
P34487 caenorhabdi
Q16787 homo sapien
P33666 escherichia
P18177 clostridium
Q92615 chlamydia p
P45508 escherichia
P23694 serratia ma
P52143 escherichia
Q9Kka3 r outer mem
P76347 escherichia

ALIGNMENTS

34 164.5 3.5 475 1 PRTG ERWCH
35 163.5 3.5 1592 1 GTF2_STRDO
36 163 3.5 918 1 YMAJ_CAEEL
37 163 3.5 1713 1 LMAJ_HUMAN
38 162.5 3.5 2003 1 YDBA_ECOLI
39 161 3.5 2366 1 TOXB_CLODI
40 160.5 3.5 1609 1 PM21_CHLPN
41 160 3.4 1250 1 YFAL_ECOLI
42 159.5 3.4 487 1 PRZN_SERMA
43 159 3.4 1569 1 YPUA_ECOLI
44 157 3.4 1655 1 OMPB_RICCN
45 156.5 3.4 2358 1 YEEJ_ECOLI

Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
Lipoprotein; Palmitate.
KW TRANSMEM 230 250 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 381 401 POTENTIAL.
FT DOMAIN 734 784 6 X REPEATS, GLY-RICH.
FT REPEAT 734 739 1.
FT REPEAT 743 748 2.
FT REPEAT 752 757 3.
FT REPEAT 761 766 4.
FT REPEAT 770 775 5.
FT REPEAT 779 784 6.
SQ SEQUENCE 953 AA; 102206 MW; 927PF56CFC884F12 CRC64;

Query Match 50.2%; Score 2332; DB 1; Length 953;
Best Local Similarity 49.1%; Pred. No. 7.1e-110;
Matches 467; Conservative 183; Mismatches 261; Indels 40; Gaps 15;

QY 1 MSINIV-IKSN---IQAGLNST-----KSGLNLYLAIPKD--YDPQKGGTLD 43
DB 5 LTNISTWKSWSLTAKSGLNRTGSLAKAGOSLTGAKIILYIPKDYQYDTEKNGLOD 64
QY 44 FIKADELIGARLAEEPNHTETAKSVDTVNQFSLTQTGTIAISATKLEKFLQKHSTNKL 103
DB 65 LVKAABELGIEVQKEGNDIAKAQTSLTGQNLVGLTERGIVLSAPOLDKLQK--TKV 121
QY 104 AKGLDSVENTDRKLGKASNVLSSTSSFLGTALAGIELDSLIIKGDADPADLAKASIDLIN 163
DB 122 QGATGSAENLTGFSNAKTVLGSIQSLGVLGMDLDEALQK-NSNELTLAKAGLELTN 180
QY 164 EILGNLSQSTOTIEAFSSOLAKLASTISQAGFNSIGNKLNQNTN-FSKTNLGLRIITGLL 222
DB 181 SLINIANSVKTLDAFGDQINQLGSLQNVNGLSLGDKLGLSGFDTSLGLDVSGLL 240
QY 223 SGISAGFALADKNASTGKVAAGFELSNQVIGNTVKAISYVLQARVAAGLSTTGAVAL 282
DB 241 SGATAALVLADKNASTRKVGAGFELANQVGNITKAVSSYILQARVAAGLSSTGPAVAL 300
QY 283 ITSSIMLAISPLAFNADKFNHANALDEFKQKPKFGYDGDHLLAEYQROVGTHIASLT 342
DB 301 IASTVSLAISPLAFAGIADKFNHAKLSLESYAERFKLGYDGDNLLAEYQRTGIDRSVT 360
QY 343 TISTALGAVSAGVAAAGVAVGTPIALLVAGVTGLSGILEASKQAMFESVANRLQCKI 402
DB 361 AINTALAAAGVSAAGRGVSIAPFIALLVSGITGVISTILQYSKQAMFHVANKHVKI 420
QY 403 LEWEKQNGQNYFDKGYDSRYAAVLANNKLFSELNKELEAERVAITQQQWNNNIGELA 462
DB 421 VEWEKNNHGNKYPFNGYDARYLANLQDNMKFLLNKLQEAERVAITQQQWNNNIGELA 480
QY 463 GITKLGRIKSGKAYADAFEDGKKVAGSNITLDAKTGIIISNSNGKTKQALLHFTSPLL 522
DB 481 GISRLGKVLGSKAYVDAFEFGKHLKADKLVLQDSANGIIDVSNSGAKTODILFRTPLL 540
QY 523 TAGTESRELNTGKYSYINKLKFRGVKNQVYTDGEASSKLPFSKVIQV-----AETE 575
DB 541 TPGDHRERVTQKYEYITKLINIRVDSWKITDGAASSTFDLTNVVQKIGIELNAGNVT 600
QY 576 GTDEIGLTVNAKAGNDIDFVGGKKNIDGGGHRVFSYKDGFGFNGITVDGTSATEAGSY 635
DB 601 KTKETKIVAKLGAGDDNVFVSGTTEIDGEGYDRVHYSR-GNYGALTIDATKETEQGSY 659
QY 636 TVNRKRVAGDIYHEVVKQETKRTETIYQRYELRKVGYGVOSTDNLKSVEVIGSQ 695
DB 660 TVNRFVETGKALHEGTSHTALVGNREKIEYR-HSNQHAGYVTKDTLKAVEEIICTS 718
QY 696 FNDVFKSGKFNDI FHSFGDILLDGGAGDRLFGKGNDRLSGDEGDLGLGGSGDVLN 755
DB 719 HNDIFKSGKFNDAFNGSGDVTIDGKGNDRLLFGKGGDDIDGNGDDFDIDGKGNLH 778
QY 756 GGAGNDVYIFRKGQNDFTLYDGTGNDKLAFADANISDIEMIERTKEGIIIVKNDHSGSINI 815
DB 779 CGKGGDDIFVHRQGDNDIITDSGNDKLSFSDSNLKDLTFFKVKHNLVI-TNSRKEKVTI 837

QY 816 PRWY-----ITSNLONYOSNKTDHKEIOLIGKGSYITSDOIDKILQDKDGTVTISOELX 871
DB 838 QDMFREADFAKEVNRKATK-DEKIEIIGQNGERITSKQVDDLI--AKNGKITQDELS 894
QY 872 KLADENKSKLSASDIASSINKLVGSMALFGTANSVSSNALQIPITOPTQGI 922
DB 895 KVVNDYELLKHS-KNVTNSLDKLSSASAFSTSSNDSNRNVLVAPTSMLDQSL 944

RESULT 3
LKTA PASSP
ID LKTA PASSP STANDARD; PRT; 947 AA.
AC P55123;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Pasteurella haemolytica-like sp. (strain 5943B).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=28165;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93239320; PubMed=8478098;
RA Chang Y.-F., Ma D.-P., Shi J., Chengappa M.M.;
RT "Molecular characterization of a leukotoxin gene from a Pasteurella
RT haemolytica-like organism, encoding a new member of the RTX toxin
RT family";
RL Infect. Immun. 61:2089-2095(1993).
CC -I- FUNCTION: Virulence factor which is cytotoxic for leukocytes but
CC is not hemolytic.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- DOMAIN: The Gly-rich region is probably involved in binding
CC calcium, which is required for target cell-binding or cytolytic
CC activity.
CC -I- DOMAIN: The three transmembrane domains are believed to be
CC involved in pore formation by the cytotoxin (By similarity).
CC -I- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC modified (By similarity).
CC -I- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L12148; AAA16444.1; -
CC InterPro; IPR001343; Hemlysn_Ca_bind.
CC InterPro; IPR003995; RtxA.
CC Pfam; PF00353; hemolysinCabin; 5.
CC Pfam; PF02382; RTX; 1.
CC PRINTS; PR00313; CABNNGRPT.
CC PRINTS; PR01488; RTXTOXINA.
CC PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
CC Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW Lipoprotein; Palmitate.
FT TRANSMEM 154 170 POTENTIAL.
FT TRANSMEM 312 333 POTENTIAL.
FT TRANSMEM 393 414 POTENTIAL.
FT DOMAIN 625 780 7 X REPEATS, GLY-RICH.
FT REPEAT 625 630 1.
FT REPEAT 730 735 2.
FT REPEAT 739 744 3.
FT REPEAT 748 753 4.
FT REPEAT 757 762 5.
FT REPEAT 766 771 6.
FT REPEAT 775 780 7.
SQ SEQUENCE 947 AA; 101559 MW; 9744F06395EF5BED CRC64;

```
Query Match          49.6%; Score 2306; DB 1; Length 947;
Best Local Similarity 50.0%; Pred. No. 1.4e-108;
Matches 466; Conservative 173; Mismatches 253; Indels 40; Gaps 16;

QY 1 MSNINV-IKSNIOAGLNST-----KSLGNLYLAIPKD--YDPQKGGTLNDFIK 46
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 4 LANSTNLKNSLSGLHKTQSLNQAGQSLKAGAKLILYIPKDYEDSGREGLQDLVK 63
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 47 AADELGIARLAEPNHTTAKKSVDTVNOFISLTQTGTIAISATKLEKPLQKHSNKLAKG 106
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 AAEDLGIEVQREERNGIATQNSLSTIQNILGFSRGGVVLSPQLDKLLQKY--KISKA 120
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 107 LDSVENIDRKLKASNVLSLSSPLGTALGIELDSLILKKGDAAPDALAKASIDLINEII 166
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 PGSSENVAKNLGNAGTLGSLGOSILGVSVMAGWDDELILKNGSELD-LAKAGLELTNSLI 179
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 167 GNLSOSTTIEAFSSQLAKLSTISQAAGFNSGNKLNQNLN-FSKTNILGLEIITGLLSGI 225
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 ENIANSVQTLDTFEEQISQGLTKONVKGSLGTGDKLKNFSGFKAGLGLEVISGLLSGA 239
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 226 SAGFALADKNASTGKKVAAGFELNQVNTKKAISSVYLAQVAAAGLSTTGAVALITS 285
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 240 TAAVLADKNASTDRKVGAGFELANOVVGNITKAVSSYILAQVAAAGLSTTGPVSALIAS 299
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 286 SIMLAISPLAFNAAADKENHANALDEFAPKPRKGYDGDHLLAEVORGVTIEASLTIS 345
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 TVALAISPLAFAGATDFNNKALAESYAEPRFKLGYEGDLSLAEYQRTGTIDASVTAVN 359
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 346 TALGAVSAGVAAAVGAVGTPIALLVAGVTGLISGLEASKQAFESVANRLOQKILEW 405
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 360 TALAISGVSAAAGSLVGPIALLVSGITGILSTILOYSQAMFEHVAHKIKVIDW 419
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 406 EKQNGQNYDPKGYDSRYAAVLANNLKFLSBLNKELEAEVIAITQQRWDNNIGELAGIT 465
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 420 EKCHGKNYFENGYSRYLADLQNMROLNKLQELQAEVIRITQQQWDDNNIGNLAGIS 479
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 466 KLERIKSKAYADAFEDKKEAGSNITLDAKTCIDISNNGKTCQALHFTSPLTAG 525
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 480 RLGEKVMGKAYADAFEEGKLKIDKTFVQLODSATSVINTSKDNNKTOHILFRPLTTPG 539
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 526 TESRELTNGKYSYINKLKFRGVKNQWQVTDGEASSKLDKFSKVIQVIA-----ETEGTD 578
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 540 VENNERIQTKYKYEITKLNINRVDKMTDGTATNSTFDLTNNVQRIEGLDHADNVTKK 599
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 579 EIGLIVNAKAGNDIDFVQGGKNNIDGGDHRVFKDGGFNITVDGTSATEAGSYTVN 638
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 600 ETKIIANLGDGNDVDFIGSGTTEVDGNGLDRLVHYSR-GDYGALTIDATNESVQGSYTVK 658
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 639 RKVARGDIYHEVKKQETKVGKRTETIOYRDELKVGKGYOSTDNLKSVVEEVIGSQFND 698
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 659 RFVETGKALHEVTATQSVLVSREKIEYR-HSNNTQHAGYITDTLKSVBEIIGTSKND 717
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 699 VFKGSKFNDIFHSGGDDLLDGGAGDRLFGKGNDRSLGDEGDDLLDGGGDDVLANGGA 758
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 718 IFKGSKFDDAFHGGGVNIGNAGNDRFLFGKGFDDIDGGGDDDFDGGGDDILHGGK 777
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 759 GNDVYIFRKGDNPTLYDGTGNDKLAFADANISDIMITERTKEGIVKRNDSHGSINIPRW 818
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 778 GNDILCTVKG-GNDSISDGGNDRLSFADSNLKLTFEKNHLMV-TNVKKEKVITQNW 835
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 819 Y-----ITSNLQYQSNKTDHKEQLIGKGYITSQDIDKILQDKDGTVITSQELKLA 874
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 836 FREADYAKTVNHYQAT-ADEKIEETIGQGRITRSKQIDELIEKGK--KIDQSELERIA 892
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 875 DENKSQKLSASDIAS-SLNKLVGSMALFGTAN 905
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 893 E--SSALLKESKFASNSLNKLVSAGAFASN 922
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
LKA3_PASHA
ID_LKA3_PASHA STANDARD; PRT; 953 AA.
```



```
Qy 109 SVENIDRLKASNVLTSLGFLCTALAGIELDSLKKGDAAPDALAKASIDLINEIGN 168
Db 127 SSESIAQNFQAKTVLISVQVSLGSLVAGWDLDEAL-QNESDQTLAKAGIELTNSLIEN 185
Qy 169 LSQSTQRTIEAFSSQLAKGSTISQAKGFSIGNKQLNL-NFSKTNLGLIEITGLLSGISA 227
Db 186 IANSVQTLDAFSEQISQFGSKLQNVKLGALGDKLKNIGGLDKAGLGLHVISGLLSGATA 245
Qy 228 GFALADKNASTGKKAAGFELSNOVIGNVTKAISSVYLAQRYAAGLSTTGAVALIITSSI 287
Db 246 ALVLADKADASTAKVAGFELANGVGNITKAVSSYILAQRVAARLSSTGPFVAALIASTV 305
Qy 288 MLAISPLAFMAADKFNHANALDEFAKQFRFGYDGDHLLAEYQGVGTIEASTTTISTA 347
Db 306 ALAISPLISFAGIADKFPRAKSLNENYAEFKLGEBSLLAEYQHGTTDASVTANTA 365
Qy 348 LGAVSAGVSAAGVAVGTPTIALIYAGVTGLISGILEASKOAMPESVANRLQGGKILEWEK 407
Db 366 LAALAGGVSAAGVAVSPIALLVSGITGVISTILQVSKQAMPEHVAANKHKNKIVEWEK 425
Qy 408 QNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAEERVITAITOQRDNNIGELAGITKL 467
Db 426 NNGGKNYFENGVDARYLANLQDNNKFLNLNKLQAEERVITAITOQDQNSNIGDLAGISRL 485
Qy 468 GERIKSGKAVADAFEDGKVKVAGSNITLDAKTGIIIDISNSNGKKTQALHFTSPLLTAGTE 527
Db 486 GERKVLGSKAVYDAFEQGHKLKADKLVLQDSAKGIIDVSNTEAKTQHILFRTPLLTGTE 545
Qy 528 SREKLTNGKYSYINKLKFGRVKNQVDTGEASSKLDPSKVIQVVA-----ETEGTDEI 580
Db 546 KRERVQTKGYEITKLNHNRVDSQIKDGAASSTFDLTNVVQRTGVDELDAENAVIKTKET 605
Qy 581 GLIYNAKAGNDIDIVGQGNKNDIGDGHDRFVSKDGGFNGNITVDGTSATAGSYTVNRK 640
Db 606 KIVATLGDGDNVVFVSGTTEIDGEGYDVRVHYSR-GNYGALTIDATKEPEQGSYTVNRF 664
Qy 641 VARGDIEVHEVVRQETKVKRKTETIQVRYDELKRVGVGYQSTDLKSVVEVIGSQFNDVF 700
Db 665 VESKALHEVTSHTALVGNREEKIEYR-HSNQOHAGYTKYDLKAVEELIGTSHNDIF 723
Qy 701 KGSKFNDIFHSGEDDLDGAGDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGN 760
Db 724 KGSKFNDIAFGDGVDTIDGNDGNDRLFGGKGDDIIDGNGDDFDIDGKGNHLLHGGKGV 783
Qy 761 DVYIFRKGNDTLYDGTGNDKLAFAANISDIMIERTKEGIIIVKRNHDSGINSIPRWY- 819
Db 784 DIFVHROGDGNDISITESGDKLSFSDSLKDLTFEKNVHHLVI-TNTKQEKVTIQNMF 842
Qy 820 ---ITSNLQYOSNKTDBKIBOLIGKDGSYITSDIQDKLDKDGTVITSQELKLADE 876
Db 843 EAEFAKTIONVATV-DDKIEEIQGGERITSQVDDLI--AKNGKIAQSELTKVVDN 899
Qy 877 NKSQKLSASDITASSLNKLGVSMALFGTANSVSSNALQPIPTQOILAPSV 927
Db 900 YOLLKYS-RDASNSLDKLISASAFSTSSND-SRNVL--ASPT-SMLDPSL 944

RESULT 5
RT2A ACTPL STANDARD; PRY; 956 AA.
AC P15377;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RTX-II toxin determinant A (APX-IIA) (Hemolysin IIA) (HLY-IIA)
DE (Cytolysin IIA) (CLY-IIA)
GN APXIIA OR CLYIIA OR HLYIIA OR APPA OR CYTC.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=715;
RN (1)
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=Serotype 5;
RX MEDLINE=90126233; PubMed=2693022;
RA Chang Y.-F., Young R., Struck D.K.;
RT "Cloning and characterization of a hemolysin gene from Actinobacillus
RL (Haemophilus) pleuropneumoniae.";
RN DNA 8:635-647 (1989).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate CUI 13261 / Serotype 9;
RX MEDLINE=92040145; PubMed=1937809;
RA Smits M.A., Braire J., Jansen H.E., Kamp E.M.,
RA Gielen A.L.;
RT "Cytolysins of Actinobacillus pleuropneumoniae serotype 9.";
RL Infect. Immun. 59:4497-4504 (1991).
CC -I- FUNCTION: One of the virulence factors of A. pleuropneumoniae,
CC which shows a weak hemolytic activity and is moderately cytotoxic
CC for alveolar macrophages and neutrophils.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- DOMAIN: The Gly-rich region is probably involved in binding
CC calcium, which is required for target cell-binding or cytolytic
CC activity (By similarity).
CC -I- DOMAIN: The three transmembrane domains are believed to be
CC involved in pore formation by the cytotoxin (By similarity).
CC -I- PTM: Palmitoylated by apxIIC. The toxin only becomes active when
CC modified (By similarity).
CC -I- SIMILARITY: BELONGS TO THE RTX PROXARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M30602; AAA87232.1; -.
DR EMBL; X61111; CAA43423.1; -.
DR PIR; B33389; B33389.
DR InterPro; IPR001343; Hemlyan_Ca_bind.
DR InterPro; IPR003995; RTXa.
DR Pfam; PF00353; hemolysincabind; 5.
DR PRINTS; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 1.
KW Hemolysin; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW Lipoprotein; Palmitate.
FT TRANSMEM 233 256
FT TRANSMEM 266 323
FT TRANSMEM 361 406
FT TRANSMEM 719 787
FT DOMAIN 719 724
FT REPEAT 719 724
FT REPEAT 728 733
FT REPEAT 737 742
FT REPEAT 746 751
FT REPEAT 755 760
FT REPEAT 764 769
FT REPEAT 773 778
FT REPEAT 782 787
FT REPEAT 787 8.
SQ SEQUENCE 956 AA; 102531 MW; BDBCABBADF14A641 CRC64;

Query Match 48.1%; Score 2236.5; DB 1; Length 956;
Best Local Similarity 47.7%; Pred. No. 4.4e-105;
Matches 453; Conservative 188; Mismatches 255; Indels 53; Gaps 15;

Qy 1 MSNINV--IKSNTQAGL-----NSTKSGLNKLYLAIPKVDYDQ 36
Db 1 MSKTLSSLSKSLQGLKNGKNKLNQAGTTLKNGLTQTHSLQNGAKKLILYIPQYDSG 60
Qy 37 KGGLTNDIFKADELGARLAEPNHTETAKSVDTVQFLSLTQTQIAISATKLEFLQ 96
Db 61 QGNGVDLVAANDLGIENVREERSNLDIAKTSPTTQTKILGFTDRGIVLFPQLDNLK 120

SEQUENCE FROM N.A.
```

```
QY 97 KXSTNKLAKGLSDVENIDRKLGKASNVLSLTSSFLGTALAGIELDLSLKKGDAAPDALAK 156
Db 121 KNP--KIGNTLGSASSISQNIQKANTVVGQISILGSLVGNLQNLQNDPQLELAK 178
QY 157 ASIDLINEIIGNLQSTQTIQAFSSQLAKLGGSTISQAQGFNIGNKLQNL--NFSKTNLGL 215
Db 179 AGLEITNELVGNIASSVQVDPAPQIQSLGSHLVKVGGLSLNKLQNLPLDKGASLGL 238
QY 216 EIITGLLSGI SAGFALADKNASTGKVAAGFELSNQVIGNTVKAISSVYLAQRAAGLST 275
Db 239 DIISGLLSGAGLTLADKEASTEKKAAGVEFANQIICNVTKAVSSYLAQRAAGLSS 298
QY 276 TGAVALITSSIMLAISPLAPNPAADKENHANALDEPAKQPKFGYDGDHLLAEYQRCVG 335
Db 299 TGPVAALITASTVALAVSPLSNVADKFKQADLIKSYSERFQKLGVDGDRLLADPHRETG 358
QY 336 TIEASLTITSTALGAVSAGVAAAAGSAVGTPIALLVAGVTGLISGLEASKQAMFESVA 395
Db 359 TIDASVTINTALAIISGVGNASAGSLVGPVALLVAGVTGLITLILEYSKQAMFEHVA 418
QY 396 NRLOQKILEWQKQNGQYFKGYDSRYAAVLANNLKPLSELNLELEAEVIAITQQRWD 455
Db 419 NKVHDRIWEVEKCH--NKNYFEQGYDSRHLADLQDNMKFLINLKELOAERVAITQQRWD 477
QY 456 NNIGELAGITKLGERIKSGKAYADAFEDGKVEAGSNITLDAKGIIDISNNGKKTQAL 515
Db 478 NOIGDLAISRDKIKSGKAYVDAFEGQHQSDYSSVQDNKNGIINISNTN--RKTQSV 536
QY 516 HFTSPLLTAGTESRRLTNGKYSYINKLKFQVKNQVQVTDGEASSKLDFSKVIOQVA--- 572
Db 537 LFRTPLLTPGEENRERIOEGKNSYITKLHIQVDSMTVTDGASSVDFTNVVQRIAVKF 596
QY 573 -----BTEGTEIGLIVNAKAGNDIIFVGQKKNIDGSDGHRVFKYSGDGFENITVD 625
Db 597 DDAGNIEESKDTK---IIANLGNAGNDNVFVSGSSTTVIDGSDGHRVHYSR--GEYCALVID 652
QY 626 GTSATEAGSYTVNRKVGADYIHEVVQEYKQETKVKRTETIOYRDELKVGKVGYSTDLN 685
Db 653 ATAETEGSYSVKRYVGSOKALHETIATHQTVGNRREEKIETR--REDDRFHTGYTVTDSL 711
QY 686 KSVEVIGSQFNDVFKSGFNDIFHSQEGDILLDGGAGDDRLFGKGNDRLSGDEGDDL 745
Db 712 KSVEEIIQSQFNDIFKSGQFDDVFGNGVDTIDGDCDDHFLFGAGDDVIDGGNGNFL 771
QY 746 DGGSGDDVLNGAGNDVYFRKGDNDTLYDGTGNDKLAFAADANISDIMEIERTKEGIIVK 805
Db 772 VGGTGNDIISGGKNDIYVHKTDGNDSTITDSGGQDKLAFSDVNLKOLTFKKVDSLEI- 830
QY 806 RNDHSGSINIPRWY----ITSNLQYQSNKTDHKLQELIGKDGSVITSDQIDKILQDKKD 861
Db 831 INQKGEKVRIGNWFLEDDLASTVANYKAT--NDRKIEEIIIGGGERITSEQVDKLI--KEG 887
QY 862 GTVITSQBELKKLADENKQKLSASDIASSLNLKVGSMALFGTANSVSSN 910
Db 888 NNQISAEALSQVNDYNTSK--DRQNVSNLAKLISSVGSFTSSSDFRNN 935

RESULT 6
LKAA_PASHA
ID_LKAA_PASHA STANDARD; PRT; 955 AA.
AC P55117;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Leukotoxin from serotype T10.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype T10;
```

```

MEDLINE=96425875; PubMed=8828217;
Lainson F.A., Murray J., Davies R.C., Donachie W.;
"Characterization of epitopes involved in the neutralization of
Pasteurella haemolytica serotype A1 leukotoxin.";
Microbiology 142:2499-2507(1996).
-!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
cell membranes and cause cell rupture by mechanisms not clearly
defined.
-!- SUBCELLULAR LOCATION: Secreted.
-!- DOMAIN: The Gly-rich region is probably involved in binding
calcium, which is required for target cell-binding or cytolytic
activity.
-!- DOMAIN: The three transmembrane domains are believed to be
involved in pore formation by the cytotoxin (BY SIMILARITY).
-!- PTM: Palmitoylated by Ikfc. The toxin only becomes active when
MODIFIED (BY SIMILARITY).
-!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; Z26247; CAAB1206.1; -.
PIR; S37145; A35254.
InterPro; IPR001343; Hemlysn_Ca_bind.
InterPro; IPR003995; RtxA.
Pfam; PF00353; hemolysinCabin; 5.
Pfam; PF02382; RTX; 1.
PRINTS; PR00313; CABNDNGRPT.
PRINTS; PR01488; RTXTOXINA.
PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
Hemolysin; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
Lipoprotein; Palmitate.
TRANSMEM 299 319 POTENTIAL.
TRANSMEM 361 381 POTENTIAL.
TRANSMEM 383 403 POTENTIAL.
DOMAIN 736 786 6 X REPEATS, GLY-RICH.
REPEAT 736 741 1.
REPEAT 745 750 2.
REPEAT 754 759 3.
REPEAT 763 768 4.
REPEAT 772 777 5.
REPEAT 781 786 6.
SEQUENCE 955 AA; 102187 MW; B60F2DB8168EBCAF CRC64;

Query Match 48.1%; Score 2236; DB 1; Length 955;
Best Local Similarity 48.2%; Pred. No. 4.7e-105;
Matches 458; Conservative 178; Mismatches 271; Indels 44; Gaps 17;

QY 5 NVIKSNITQA---GLN-----STKSLKLNLYLAIPKD--YDPQKGGTLNDFIKAA 48
Db 12 NGIRSTLTATGGLNRAGQSITQAGQTLKNGAKKIIYIPKDYKDYSGSGLQDLVKAA 71
QY 49 DELGIARLASEPNHITETAKSVDTVNOFLSTQFGIAISATKLEKFLQKSTNKLAKGLD 108
Db 72 EELGIEVQKEGNDIAKAQTSIGTIQNVLGITGERGIVLSAPQLDKLQK--NKGVGQALG 128
QY 109 SVENIDRKLGKASNVLSLTSSFLGTALAGIELDLSLKKGDAAPDALAKSIDLINEIGN 168
Db 129 SSESTAQNFSQAQKTVLSGVQGNSTVLAGMDLDEAL--QNESDQUTLAKAGLELNTSLIEN 187
QY 169 LSQSTQTIQAFSSQLAKLGGSTISQAQGFNIGNKLQNL--NFSKTNLGLIITGLLSGISA 227
Db 188 IANSVQILDAFSEQISQFGSKLQNVKGLGALGDKLKNIGGLDKAGLGLDKVSRLLSGATA 247
QY 228 GFALADKNASTGKVAAGFELSNQVIGNTVKAISSVYLAQRAAGLSTTGGAVALITSSI 287
Db 248 ALVLADKDASTAKKVGAGFELANQVGNITKAVSSYLAQRAAGLSTGTPVAALIASTV 307
```


QY 403 LEWKGONGQYDPKGYDSRYAAVLANNKFLSELNKELEAERVIAITQORWNNIGELA 462
 Db 437 DEWEKKY-GKNYFENGDAHKAFLDSFLLSNFKQYETERAVLITQORWDEYIGELA 495
 QY 463 GITLGERIKSGAYADAFEDGKKVEAG----SNITLDKATGIIDISNSGKKTQALHFT 518
 Db 496 GITCKGDKLSSGAYVDFQEGKLEKPPDDFSKVFPDPTKGEIDISNS--QTSTLLKFV 553
 QY 519 SPLTAGTESRERTNGKYSYINKLFGVRVKNQVDTG--EASSKLFDFSVIQR----- 570
 Db 554 TPLTPGTSRERTQTGKYEYITKLVVVGKDKW--VNVGKDKGAVYDYNLIQHAHSSS 612
 QY 571 VAETEGTDEGLIVNAGANDDI FVGCGKMDICDGHDRVFSKGGGFGNITVDGTSAT 630
 Db 613 VARGEYREVLVSHLGNLNDKVFVLAAGSAEIHAGEHDVVYDKT--DTGLLVIDGTYAT 671
 QY 631 EAGSYTVNRKARG--DIYHEVYKRGQETKVGKRTETIQYRDYELRKVG--YGVQSTDNLSKV 688
 Db 672 EQGRYSVTRELSGATKILREVINKQSAVGKREBTEYRDYELTQSGNSNLKAHDELHSV 731
 QY 689 BEVIGSQDFVFKSKNDIFHSBEGDLDLGGAGDRFLFGKGNDRLSGDEGD----- 742
 Db 732 EEIIGSNQORDEFKSGKFRDIPHGADGDDLLNGNDGDDILYDGNLGRGNDGNDQLYGG 791
 QY 743 -----DLLDGGSGDDV 753
 Db 792 EGNDKLLGGNNYLSGGDGNDELQVNGFVLRGGKDDKLYSGSGSDDLDDGEGNDY 851
 QY 754 LGGAGNDVYIFRKGNDGNTLYD---GTGNDKLAAPADANISDIERTKGIIVKRNDSH 810
 Db 852 LEGDGSDFVYRSTSGNHTYDQKSDLDKLYSDFSPRLLVKGVDDNLVLRNESS 911
 QY 811 ---GSINPRWYTSNLQVSNKTDHKEBQICDGSYITSDDIKLQDKDGTVITS 867
 Db 912 HNNGLVLTIKMP-----KEGNKYNHKEIQIVDKNGRKLTAENLGTTFKNAPKADNLN 964
 QY 868 QELKKLADENKSKLASDIASSLNKLGVSMALFGTA 904
 Db 965 YATKE--DQNES---NLSSLKTELKSIITNAGNEGVA 996

RESULT 9
 ID RT31 ACTPL STANDARD; PRT; 1049 AA.
 AC P55130;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE RTX-III toxin determinant A from serotype 2 (APX-IIIa) (Cytolysin
 DE IITA) (CLY-IIIa)
 GN APXIIIA OR CLYIIIA OR RTXA OR PTXA.
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Actinobacillus.
 OX NCBI_TaxID=715;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Serotype 2;
 RX MEDLINE=93263992; PubMed=8494611;
 RA Chang Y.-P., Shi J., Ma D.-P., Shin S.J., Lein D.H.;
 RT "Molecular analysis of the Actinobacillus pleuropneumoniae RTX
 RT toxin-III gene cluster";
 RL DNA Cell Biol. 12:351-362 (1993).
 RN [2]
 RP SEQUENCE OF 828-1049 FROM N.A.
 RC STRAIN=1536 / Serotype 2;
 RX MEDLINE=95012630; PubMed=7927703;
 RA Jansen R., Briare J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J.,
 RA Smits M.A.;
 RT "Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (ApX)
 RT operons: characterization of the ApXIII operons";
 RL Infect. Immun. 62:4411-4418 (1994).
 CC -!- FUNCTION: Does not have hemolytic activity but shows a strong

CC cytotoxicity towards alveolar macrophages and neutrophils.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- DOMAIN: The Gly-rich region is probably involved in binding
 CC calcium, which is required for target cell-binding or cytolytic
 CC activity (By similarity).
 CC -!- DOMAIN: The three transmembrane domains are believed to be
 CC involved in pore formation by the cytotoxin (By similarity).
 CC -!- PTM: Palmitoylated by apxIIIC. The toxin only becomes active when
 CC modified (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L12145; AAA21924.1; -;
 CC EMBL: X80056; CAB37652.1; ALT_SEQ.
 CC FIR: S51784; S51784.
 CC InterPro: IPR001343; Hemlysn_Ca_bind.
 CC InterPro: IPR003995; RtxA.
 CC Pfam: PF00353; hemolysinCabin; 6.
 CC Pfam: PF02382; RTX; 1.
 CC PRINTS: PRO0313; CABNDNGRPT.
 CC PRINTS: PRO1486; RYTXOXINA.
 CC PROSITE: PS00330; HEMOLYSIN CALCIUM; 3.
 CC Toxin; Cytolysis; Repeat; Calcium; Transmembrane; Lipoprotein;
 CC Palmitate.
 FT TRANSMEM 154 170 POTENTIAL.
 FT TRANSMEM 315 331 POTENTIAL.
 FT TRANSMEM 397 413 POTENTIAL.
 FT DOMAIN 753 858 11 X REPEATS, GLY-RICH.
 FT REPEAT 753 758 1.
 FT REPEAT 762 767 2.
 FT REPEAT 771 776 3.
 FT REPEAT 780 785 4.
 FT REPEAT 789 794 5.
 FT REPEAT 798 803 6.
 FT REPEAT 807 812 7.
 FT REPEAT 826 831 8.
 FT REPEAT 835 840 9.
 FT REPEAT 844 849 10.
 FT REPEAT 853 858 11.
 SQ SEQUENCE 1049 AA; 112491 MW; F99846BFD4ESCE72 CRC64;
 Query Match 41.4%; Score 1924; DB 1; Length 1049;
 Best Local Similarity 43.0%; Pred. No. 2.5e-89;
 Matches 428; Conservative 152; Mismatches 296; Indels 120; Gaps 21;
 QY 8 KSNIQAGLNSKSGLK-----NLYLAIPKDYDPQKGTLDNFI 45
 Db 19 KRQVKGYDVTGNGLVGSQAKLQALAAKAVQKGNKLVLPKEYDGSVGNPFDLV 78
 QY 46 KAADGLIARLAEPNHTETAKSVDTVQFLSITQTGIAISATKLEKFKQKSTNKLAK 105
 Db 79 KAABELGIQVYVNRNELEVAHKSGLTADQFLGTERTGLTFAPQLDQFLQKHSKSNV 138
 QY 106 GLDSVENIDRKLKASNVLSLTSFLGTALAGLELDSLIKKGDAAPDALAKASIDLNEI 165
 Db 139 GSSTGDAVS-KLAKSQTIIISGIVSLTGLVAGINLNEAIISSGSELE-LAEAGVSLASEL 196
 QY 166 IGNLSOSTQTEAFSSQLAKGSTI SOAKGFSINGNKQLN---FSKTNLGLIEITGLL 222
 Db 197 VSNIAKGTITIDAFITQIQFNGKLAENAKGLGVGROLQNGISGALSKTGLGLDISSL 256
 QY 223 SGISAGFALADKNASTCKKVAAGFELSNOVINVTKAISSVLAQRAAGLSTGVAAL 282
 Db 257 SGVTRSPALRNKNASTSTKVAAGFELSNOVINVTKAISSVLAQRAAGLSTGVAAL 316
 QY 283 ITSSIMLAISPLAPMAADKFNHANALDEPAKFRKFGYDGDHLLAEYQRGVGTIEASLT 342

```

Db 317 IASSISLAISPLAFURVADNFRNKEICEFAERFKLGYDGDGKLLSEYHEAGTIDASIT 376
Qy 343 TISTALGAVSAGVSAAGVGTPIALLVAGVTGLISGLEASQAMFESVANRLQOKI 402
Db 377 TISTALSIAAGTAASAGALVGAPITLLVTGITGLISGLEAFKQPMLDHVASKIGNKI 436
Qy 403 LEWEKONGQNYFDKGYDSRVAAYLANNLKPLSELNKELEAERVIAITQOBDNNGIGELA 462
Db 437 DEWEKKY-GKNYFENGYDARHAKFLEDSFLLSNFKQYETERAVLLITQOBDNNGIGELA 495
Qy 463 GITKLGERIKSGKAVADAFDCKKVEAG-SSNITLDAKTGIIDISNSCKKTOALHFT 518
Db 496 GITGKGDLSGSKAVYDFQEGKLEKDPDFSKVDFPTKGEIDISNS--QTSTLLKFV 553
Qy 519 SPLTAGTESRERLTNGKYSYINKLFGVRVKNQVTDG--BASSKLDPSKVIQR----- 570
Db 554 TPLLTGPGTESRERTGTGKEYITKLIVKGDKW-VVNGVKDKGAVDYDTNLIQAHISSS 612
Qy 571 VAETEGTDEIGLIVNAKAGNDIFVGGQKMMIDGGDGHDRVPYKDGFGNITVDGTSAT 630
Db 613 VARGEYREVLVSHLGNNGNDKVFLAAGSABEIHAGEGHDVVVYDKT-DTGLLVIDGTAT 671
Qy 631 EAGSVTVNRKVARG-DIVHEVVKQETKVGKRTETIOVRDYELRKVG-YGYOSTDNLKSV 688
Db 672 EGQYSVTVRELSGATKILREVKNQKAVAGKEETLEYRDVYELTQSGNSNLKAHDELHSV 731
Qy 689 EEVIGSQDNFVFKSGKFNDFIHSBGDDLLDGGAGDRLFGKGNDRLSGD----- 739
Db 732 EE-IGSNQRDEFKSGKPRDIFHGADGDDLLNGNDGDDILYDGKGNDELGRGNDQLYGG 790
Qy 740 EGD-----DILDGGSGDDV 753
Db 791 EGDDKLLGGNGNYLSSGDDNDELQVLGNGFNVLRGKGGDKLYGSSGSDLLDGGEGNDY 850
Qy 754 LGGAGNDVYIFRKGDDGNDTLVD---GTGNDKLAFADANISDIMERTEKGIIVK--RND 808
Db 851 LEGDGSDFYVYRSTSGNHTIYDQKASDSKLYLSDSLSPNVLKRVNDNLEFRSNNS 910
Qy 809 HSGSINIPRWYITSNLQYQSNKTDHKEIQIGKDGYSVITSQIDKILQDKKDGTVITSQ 868
Db 911 NSGVLTIKDWFPGKGSYN-----HKIEQIVDKNGRKLTAGNLGNPHD---TQASS 959
Qy 869 ELKLDENKSKLSASDIASLKLKVGSMALFGTA 904
Db 960 LLKNVTQBNESNLSS--LKTELKGIITNAGNFGVA 993

RESULT 10
HLVA_ECOLI
ID HLVA_ECOLI STANDARD; PRT; 1024 AA.
AC P08715;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hemolysin, plasmid.
GN HLVA.
OS Escherichia coli.
OG Plasmid Inc12 pHLV152.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Hess J., Wells W., Vogel M., Goebel W.;
RT "Nucleotide sequence of a plasmid-encoded hemolysin determinant and
its comparison with a corresponding chromosomal hemolysin sequence.";
RL FEMS Microbiol. Lett., 34:1-11 (1986).
RN [2]
RP PALMITOYLATION OF LYS-564 AND LYS-690.
RX MEDLINE=95099325; PubMed=7801126;
RA Stanley P., Packman L.C., Koronakis V., Hughes C.;
RT "Fatty acylation of two internal lysine residues required for the

```

```

RT toxic activity of Escherichia coli hemolysin.";
RL Science 266:1992-1996 (1994).
RN [3]
RP PALMITOYLATION OF LYS-564 AND LYS-690.
RX MEDLINE=96404790; PubMed=8808931;
RA Ludwig A., Garcia F., Bajer S., Jarchau T., Benz R., Hoppe J.,
Goebel W.;
RA "Analysis of the in vivo activation of hemolysin (HlyA) from
Escherichia coli.";
RT J. Bacteriol. 178:5422-5430 (1996).
RL J. FUNCTION: Bacterial hemolysins are exotoxins that attack blood
cell membranes and cause cell rupture by mechanisms not clearly
defined.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DOMAIN: The gly-rich region is probably involved in binding
calcium, which is required for target cell-binding or cytolytic
activity.
CC -1- DOMAIN: The three transmembrane domains are believed to be
involved in pore formation by the cytotoxin.
CC -1- PTM: palmitoylated by hlyC. The toxin only becomes active when
modified.
CC -1- MISCELLANEOUS: The hemolysin of E. coli is produced predominantly
by strains causing extraintestinal infections, such as those of
the urinary tract.
CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M14107; AAA98233.1; -.
DR InterPro; IPR001343; Hemolysn_Ca_bind.
DR InterPro; IPR003995; RTXa.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 4.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 4.
KW Hemolysin; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
Lipoprotein; Palmitate; Plasmid.
FT TRANSMEM 238 260 POTENTIAL.
FT TRANSMEM 268 327 POTENTIAL.
FT TRANSMEM 365 411 POTENTIAL.
FT DOMAIN 724 870 16 X REPEATS, GLY-RICH.
FT REPEAT 724 729 1.
FT REPEAT 733 738 2.
FT REPEAT 742 747 3.
FT REPEAT 751 756 4.
FT REPEAT 760 765 5.
FT REPEAT 769 774 6.
FT REPEAT 778 783 7.
FT REPEAT 787 792 8.
FT REPEAT 796 801 9.
FT REPEAT 807 812 10.
FT REPEAT 817 822 11.
FT REPEAT 826 831 12.
FT REPEAT 835 840 13.
FT REPEAT 844 849 14.
FT REPEAT 856 861 15.
FT REPEAT 865 870 16.
FT LIPID 564 564 PALMITATE.
FT LIPID 690 690 PALMITATE.
SQ SEQUENCE 1024 AA; 110201 MW; 83944917F76C945B CRC64;

Query Match 39.8%; Score 1848.5; DB 1; Length 1024;
Best Local Similarity 41.7%; Pred. No. 1.5e-85;
Matches 418; Conservative 156; Mismatches 312; Indels 117; Gaps 22;
Qy 8 KSNITQAGLNTSKSLKLNLYLAIPKDYDPQKGGTLNDFIKADELGIARLAEPNHTAK 67

```


Db 32 KDALKKAAEOTRAGNRLILIPDKYKG-GSSLDLVRTADELGIEVQDERNGTAITK 90
Qy 68 KSVDTVNOFSLTGTGIAISATKLEKFLQKH--STNKLAKGLDSVENIDRKLKASNVLS 125
Db 91 QVFGTAELKGLTGERVTITAPQLDKLQYKAGNLTGGG---AENIGNGLKAGGILS 147
Qy 126 TLSSFLGTALAGIELSLIKK---GDAAPDAKAKASIDLINEIIGNLSQSTQITAEFS 181
Db 148 TFQPLGTALSSMKIDELIKKQSGNVSSSELAKSIELINQLVDTVASLNNVNSFSQ 207
Qy 182 QLAKLGTSTIOAKGFSNIGKNLQNL-NFSKTNLGLLEITIGLISGISAFGLADKNASTGK 240
Db 208 QNLTLGSLVNTKHLGVGNKQLNPLNDINAGLDTVSGILSAISAFILSNADADTRT 267
Qy 241 KVAAGFELSNOVIGNVTKATISSYVLAORVAAGLSTTGAAVALITSSIMLAISPLAFNAA 300
Db 268 KAAAGVELTTLKVLGNVGKISQYIIAQRAGQGLSTSAAGLTIASAVTLAISPLSLSIA 327
Qy 301 DKFNHALDEFKQFRKFGVGDGPHLAAYVORGVTIEASLTITSTALGAVSAGVSAAV 360
Db 328 DKFRANKIEYSQRFKGLGVGDSLLAAFHKETGAIDASLTITSTVLASVSSGISAAAT 387
Qy 361 GSAVGTPIALVAGVTGLISGILEASQAMPESVANRLQGLKEWKGQNGQNYFDKGYD 420
Db 388 TSLVGAIPVSLVAGVTGLISGILEASQAMPESVANRLQGLKEWKGQNGQNYFDKGYD 446
Qy 421 SRYAAYLANNLKPLSELNKELEABERVIAITQORWNNIGELAGITKLERIKSGAYADA 480
Db 447 ARHAAFLDNPKLISQYNKEYSVERSVLITQOHMDTLIGELAGVTRNGDLSKSYIDY 506
Qy 481 PEDGKVEAGSN---ITLDAKTGIDISNGKKTOALLHTSPILLTAGTRESRELTNGK 536
Db 507 YEBGKLEKKKDEFOKQVDFPKGNIDLSDS--KSSTLLKFPVPLLTGPEIRERRSGK 564
Qy 537 YSYINKLKFRGVKNQVTD-GEASSKLDFFSVIQRVAETEGT-DEIGLIVNAKAGNDIIF 594
Db 565 YEYITELLVKGDKVTKGVQDKAVDYNSNLIOHASVGNNOYREIRIESHLGDGDDKVP 624
Qy 595 VQGKMNIDGGDHRVFPYSK-DGGFNGITVDGTSATEAGSYTVNRKVARGD---IYHEVV 651
Db 625 LSAKSANIYAGKHGHDVVYDKTDGY--LTIDGTKATEAGNYTVTR-VLGGDVKVLQEVV 681
Qy 652 KROETKVKRTTTOYRDYELRKV-GYCYQSTDNLSVEEVIGSQFNDVPKSGKFNDIEH 710
Db 682 KEQEVSVGKRTKTYRSYEFTHINGKNLTETDNLYSVEELIGTTRADKFGSKFTDIEH 741
Qy 711 SGEGLDLDGAGDDR----- 726
Db 742 GADGDDLEIENGNDRLYGDGNDQLYGGDNDKLVAGNNYLNNGDGD 801
Qy 727 -----LPGKGNDRSLDGGDDLLDGGSGDDVNLGAGNDVYIFRKGDNNDT 773
Db 802 EFQVQGSNLAKNVLFGGKNDKLYGSEGADLLDGGEGDLDKGGYNDIYRLSYGCGHHI 861
Qy 774 LYDGTG-NDKLAPADANISDMIRTEKGIIVKND-----HSGSINTPRWITSNLQ 825
Db 862 IDDDGKEDKLSLADIDFRDVAFRKGNNDLIMYKGBGNVLISGHKNGITFRNWF-----E 916
Qy 826 NYQSNKTDHKEQLGKDGSVITSDIDKILQDKK-----DGTVTTSQ----- 868
Db 917 KESGDINHEIEQIFDKSGRIITPDSLKKALEYQORNNKASYVYNDALAYGSGDNLNPL 976
Qy 869 ---ELKKLADENKSQKLSASDIASSLNKLGVSMALFGTA-NSVS 908
Db 977 INEISKIISAAGSFDVKEERTAASLLQLSGNASDFSYGRNSIT 1019

RESULT 11
ID HLVI ECOLI
AC P09983
DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hemolysin, chromosomal.
GN Escherichia coli.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J96 / Serotype O4;
RX MEDLINE=85234404; PubMed=3891743;
RA Feilme T., Pellett S., Welch R.A.;
RT "Nucleotide sequence of an Escherichia coli chromosomal hemolysin.";
RL J. Bacteriol. 163:94-105 (1985).
RN [2]
RP SEQUENCE OF 1-44 FROM N.A.
RC STRAIN=2001;
RX MEDLINE=85258115; PubMed=3894051;
RA Nicaud J.-M., Mackman N., Gray L., Holland I.B.;
RT "Characterisation of HlyC and mechanism of activation and secretion
of haemolysin from E. coli 2001.";
RL FEBS Lett. 187:339-344 (1985).
CC -1- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
cell membranes and cause cell rupture by mechanisms not clearly
defined.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DOMAIN: The Gly-rich region is probably involved in binding
calcium, which is required for target cell-binding or cytolytic
activity.
CC -1- DOMAIN: The three transmembrane domains are believed to be
involved in pore formation by the cytotoxin.
CC -1- PTM: Palmitoylated by hlyC. The toxin only becomes active when
modified.
CC -1- MISCELLANEOUS: The hemolysin of E.coli is produced predominantly
by strains causing extraintestinal infections, such as those of
the urinary tract.
CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL; M10133; AAA23975.1; -;
DR EMBL; X02768; CAA26546.1; -;
DR PIR; A24433; LEECA
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCbind; 6.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 4.
KW Hemolysin; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW Lipoprotein; Palmitate.
FT TRANSMEM 237 259 POTENTIAL.
FT TRANSMEM 267 326 POTENTIAL.
FT TRANSMEM 364 410 POTENTIAL.
FT DOMAIN 723 869 16 X REPEATS, GLY-RICH.
FT REPEAT 732 737 1.
FT REPEAT 741 746 2.
FT REPEAT 750 755 3.
FT REPEAT 759 764 4.
FT REPEAT 768 773 5.
FT REPEAT 777 782 6.
FT REPEAT 786 791 7.
FT REPEAT 795 800 8.
FT REPEAT 806 812 9.
FT REPEAT 812 812 10.


```
DR EMBL; D16582; BAA04014.1; -.
DR EMBL; U04954; AAB17220.1; -.
DR EMBL; X73116; CAA51546.1; -.
DR PIR; I39641; I39641.
DR InterPro; IPR001343; Hemlyen_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 6.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXOKINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 2.
KW Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW Lipoprotein; Palmitate.
FT TRANSMEM 226 256 POTENTIAL.
FT TRANSMEM 297 326 POTENTIAL.
FT TRANSMEM 367 406 POTENTIAL.
FT DOMAIN 722 845 13 X REPEATS, GLY-RICH.
FT REPEAT 722 727 1.
FT REPEAT 731 736 2.
FT REPEAT 740 745 3.
FT REPEAT 749 754 4.
FT REPEAT 758 763 5.
FT REPEAT 767 772 6.
FT REPEAT 776 781 7.
FT REPEAT 785 790 8.
FT REPEAT 794 799 9.
FT REPEAT 813 818 10.
FT REPEAT 822 827 11.
FT REPEAT 831 836 12.
FT REPEAT 840 845 13.
FT CONFLICT 210 217 AMPYLTLA -> GNALSNTR (IN REF. 2).
FT CONFLICT 581 581 E -> Q (IN REF. 2).
FT CONFLICT 687 688 TC -> R (IN REF. 2).
FT CONFLICT 1015 1015 F -> L (IN REF. 2).
SQ SEQUENCE 1023 AA; 110129 MW; 183C7C15EE57DB55 CRC64;

Query Match 37.2%; Score 1729.5; DB 1; Length 1023;
Best Local Similarity 39.5%; Pred. No. 1.4e-79;
Matches 403; Conservative 177; Mismatches 312; Indels 127; Gaps 28;

Qy 8 KSNITQAGLNSTKSGL-----KNLYLAIPKDYDPQKGGTINDFIKADELGIARLAEEP 60
Db 21 KSAASKGAGALKNGLGQVKQAGKLLVYPKDYQASTGSSLDNLVKAALGIEVHRSEK 80
Qy 61 NHTETAKSVDTNQFLSLTGTIAISATKLEKFLQKSTNKLAKGL-DSVENIDRLKLG 119
Db 81 NGTALAKELFOTTEKLLGFSERGIALFAPQDKLLNKN--QKLSKSLGSGSEALGQRLNK 138
Qy 120 ASNVLSLTSSPLGTALAGIELSLIKK----GDAAPALAKASIDLINEIIGNLSQSTQT 175
Db 139 TQTALSALQSFLGTAGMDLSDLLRRNRGDEYSGSELAKAGVDLAAQLVDNITASATG 198
Qy 176 IEAFSSQLAKGS---TISQAKGSNTGNKLNQNF-SKTNLWGLEIITGLLSGTSAGFAL 231
Db 199 VDAPAEQLGLAMPVLTIA----LSGLASKLNNLPDLISLAGPGFDVSGILSVVSASFIL 254
Qy 232 ADKNASTGKKVAAGFELSNGVIGNVTKAISSVLAORVAAGLSITGVAALITSSIMAI 291
Db 255 SNKDADAGTKAAAGIEISTKILGNIGKAVSQYIIAORVAAGLSITTAATGGLIGSVVALAI 314
Qy 292 SPLAFMAADAFNANALDEFKFRKFGYDGDHLLAEYQSGVGTIEASLTITITGALCAV 351
Db 315 SPLFLNVADKPERAKQLEQYSEKFRKFGYDGDHLLAEYQSGVGTIEASLTITITGALCAV 374
Qy 352 SAGVSAAGVAVGTPITALLVAGVTGLTISGLEASQKAMFESVANRLQKILEWONGG 411
Db 375 SAGVGAATGSLVGPVAAVSAITGIISGILDASKQAIFERVATKLANKIDEWEKKG-G 433
Qy 412 QNYFDKGVDSYAYLANNKFLSELNKELEBAERVIAITQORWNNNIGELAGITKLGRI 471
Db 434 KNYFENGVDARHSFALEDTFFELSSQYNKEYSVERVAITQORWVDNIGELAGITRKGA 493
Qy 472 KSGKAYADAFEDGKKVEAG-----SNITLDKTIIDISNSNGKKTQALHFTSPLLTAGTE 527
```

```
Db 494 KSKKAYVDPEEGKLEKOPDRFDKVKVDFPLEGKIDLSSIN--KTTLLKFIITPVFTAGEE 551
Qy 528 SRRLTNGKYSYINKLKFRGVQWQVTDGDEASSKL--DFSQVIOVARTETGDEIGLIVNA 586
Db 552 IREKQTKGYEYMTFLVKGEKXWVTVGVESHNAIYDYNLIQLAIDKKG-EKRQVTVIES 610
Qy 587 KAG--NDDIFVGOGKWNIDGGDGHDRVFSK-DGGFNGITVDGTSTATEAGSYTNRKV-A 642
Db 611 HLGEKNDRIYLLSSGSSIVVAGNCHDVAYDKTGTG--LTFDQSAQKAGEYIVTKELKA 668
Qy 643 RGDYHWEVKRQETKVKGR-TETIQYRDVELR--KVGYGQSTDNLSKVEEIVGSGQFNDV 699
Db 669 DVKLVKEVVKTDISVCKTCEKLYRDVELSPFELGNGIRAKDELHLSVEEIIIGSNRKDK 728
Qy 700 FKSGKNDIFHSGEGD-----DILLDGGAGDRLFGGKNDRLSGDEG 741
Db 729 FFGSRETFDIFHGAKGDDIEYNGDGHILDYGGDGDNDLHGGNDLHLYGGNDRLIGGKG 788
Qy 742 DDLDDGGSGDD-----VLNGGAGNDVY-----IFRKGDDNDTLYDGTGND--- 781
Db 789 NNFLEGGDGDDELQVFEQYNNVLLGGAGNDILYSGDGTNLFDCGVGNDKLYGSLGKDIYR 848
Qy 782 -----KLAFADANI-----SDIMIERTKEGIIIVKRNDSHG-SI 813
Db 849 YSKEYGRHIIIEKGGDDTLLSLDSFKDVGFRIGDILLVNRKIRGTYLYHEDYNGNAL 908
Qy 814 NIPRWITSNLQYQSNKTDHKEIQLEKDGSIYTSQIDIKLQDKDGTGTVITSQELKKL 873
Db 909 TIKDMF--KEGKEGQNN---KIEKIVKDGAYVLQYLTETAPGTAPGNGYFNGLEBKLY 962
Qy 874 ADENKSKQ----KLSASDIASSLNKLKVS-----MALFGTANSVSSNALQPIQTQ 917
Db 963 YGEGYNALPOLRKDIEQIISSTGALTGEHGVLVGAGGPLAYSNSPNSIPNFSNYLTQ 1021

RESULT 13
RT11 ACTPL
ID RT11_ACTPL STANDARD; PRT; 1023 AA.
AC P55128;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RTX-I toxin determinant A from serotypes 1/9 (APX-IA) (Hemolysin IA)
DE (HLY-IA) (Cytolysin IA) (CLY-IA).
DE APXIA OR CLYIA OR HLYIA.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=715;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S 4074 / Serotype 1;
RX MEDLINE=91348845; PubMed=1879928;
RA Frey J., Meier R., Gygi D., Nicolet J.;
RT "Nucleotide sequence of the hemolysin I gene from Actinobacillus
RT pleuropneumoniae."
RL Infect. Immun. 59:3026-3032 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S 4074 / Serotype 1;
RX MEDLINE=94237497; PubMed=8181764;
RA Frey J., Haidmann A., Nicolet J., Boffini A., Prentki P.;
RT "Sequence analysis and transcription of the apxi operon (hemolysin I)
RT from Actinobacillus pleuropneumoniae."
RL Gene 142:97-102 (1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate CVI 13261 / Serotype 9;
RX MEDLINE=93366425; PubMed=8359891;
RA Jansen R., Briaire J., Kamp E.M., Gielkens A.L.J., Smits M.A.;
RT "Structural analysis of the Actinobacillus pleuropneumoniae-RTX-toxin
RT I (Apxi) operon.";
```


28-FEB-2003 (Rel. 41, Last annotation update)
Leukotoxin.
LTKA OR LTA.
OS Actinobacillus actinomycetemcomitans (Haemophilus
actinomycetemcomitans).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OC NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JP2;
RC MEDLINE=89359382; PubMed=2670940;
RA Lally E.T., Golub E.E., Kieba I.R., Taichman N.S., Rosenbloom J.,
RA Rosenbloom J.C., Gibson C.W., Demuth D.R.;
RT "Analysis of the Actinobacillus actinomycetemcomitans leukotoxin
gene. Delineation of unique features and comparison to homologous
toxins.";
RT
RT
RRL J. Biol. Chem. 264:15451-15456(1989).
CC -I- FUNCTION: One of the virulence factors of A.actinomycetemcomitans
CC might be a cytotoxin, possibly the membrane-bound hemolysin.
CC -I- SUBCELLULAR LOCATION: Outer-membrane associated or secreted (By
CC similarity).
CC -I- DOMAIN: The Gly-rich region is probably involved in binding
CC calcium, which is required for target cell-binding or cytolytic
CC activity.
CC -I- DOMAIN: The three transmembrane domains are believed to be
CC involved in pore formation by the cytotoxin (BY SIMILARITY).
CC -I- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC modified (By similarity).
CC -I- MISCELLANEOUS: Its target cell specificity is restricted to human
CC and some non-human cells of the monomyelocytic lineage.
CC -I- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; M27399; AAA21922.1; --
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 7.
DR Pfam; PF02382; RTX; 1
DR PRINTS; PR00313; CABNNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 5.
DR Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW Lipoprotein; Palmitate.
FT TRANSMEM 339 359 POTENTIAL.
FT TRANSMEM 408 429 POTENTIAL.
FT TRANSMEM 477 501 POTENTIAL.
FT DOMAIN 722 844 14 X REPEATS, GLY-RICH.
FT REPEAT 722 727 1.
FT REPEAT 731 736 2.
FT REPEAT 740 745 3.
FT REPEAT 749 754 4.
FT REPEAT 758 763 5.
FT REPEAT 767 772 6.
FT REPEAT 776 781 7.
FT REPEAT 785 790 8.
FT REPEAT 794 799 9.
FT REPEAT 803 808 10.
FT REPEAT 812 817 11.
FT REPEAT 821 826 12.
FT REPEAT 830 835 13.
FT REPEAT 839 844 14.
SQ SEQUENCE 1050 AA; 114194 MW; 38DF9AA24649F662 CRC64;
Query Match 36.9%; Score 1712.5; DB 1; Length 1050;
Best Local Similarity 41.2%; Pred. No. 1e-78;

PT REPEAT 1556 1561 23.
FT REPEAT 1565 1570 24.
FT REPEAT 1574 1579 25.
FT REPEAT 1583 1588 26.
FT REPEAT 1593 1598 27.
FT REPEAT 1605 1610 28.
FT LIPID 860 860 PALMITATE.
FT LIPID 983 983 PALMITATE.
FT MUTAGEN 188 188 D->E,N,Y,H: LOSS OF ACTIVITY.
FT MUTAGEN 190 190 D->N,Y,H: LOSS OF ACTIVITY.
FT MUTAGEN 298 298 H->R,P,L: LOSS OF ACTIVITY.
FT MUTAGEN 301 301 E->Q,K: LOSS OF ACTIVITY.
SQ SEQUENCE 1706 AA; 177506 MW; F00744524BDD442E CRC64;

Query Match 17.3%; Score 802; DB 1; Length 1706;
Best Local Similarity 27.1%; Pred. No. 1.1e-32;
Matches 246; Conservative 142; Mismatches 347; Indels 174; Gaps 26;

Qy 106 GLDSVENIDRKLGKASNVLSLTSLFGLTALAGIELDSLI-----KKGDAAP----- 151
Db 403 GYDSLDGV-----GSRFSLEGVSDMAVEAAELEMTQVHLHAGARQDDAEPGVSGSAHW 458
Qy 152 -----DALAKA-----SIDLINEI-----IGNLSQSTQTIAPFSSQLAKLSTISQ-A 193
Db 459 GORALQGAQAAVAAQRLVHAIFALMTQFGRAGSTTNPQEAASLSAAVFLGGEASSAVAETV 518
Qy 194 KGFSGNIGKQNLNFSKTNLGLLEIITGLLSISAGFALADKNASTGKVAAGFELSNOVI 253
Db 519 SGFPR-GSSRWAGGFGVAGGAMALGGGTAAAVGAGMSLTD-DAPAGQKAAAGABIALQLT 576
Qy 254 GNVTKATSSVYL-----AORVAAGLSTTGAAVLAITSSIMLAISPLAFNNAADKENHANAL 309
Db 577 GGTVELASSIALAALAAARGVTSGLQVAGASAGAAGALAAALSPMEIYGLVQQSHYADQL 636
Qy 310 DEFAKQFRKFGYDGHLLAEYQVRGVTIEASLTITISTALGAVSAGVSAVAAVGSVAGTPIA 369
Db 637 DKLAQESSAYGEGDALLAQLYRDKTAEGAVAGVSAVLSVTVGAASVIAAASVVGAPVA 696
Qy 370 LLVAGVTGLSIGILEASKQAMFESVANRLOKILIEWEKONGQYFKGYDSRVAAYLAN 429
Db 697 VVTSLLTGALNGILRVQVQPIIEKLANDYARKI---DELGGPQYAFKNIQARH-EQLAN 752
Qy 430 N---LKFSELNKELEAEVIAITQORWNNIGLAGITTKLGERIKGKAVADAFEDGKK 486
Db 753 SDGLRKLADLOAGWNAASSVLGVQTTSEISKSALELAITGNADNLKSVDFVDFVQGER 812
Qy 487 VEAGSNITLDAKTIIDISNKGKTKQALHTSPLLTAGTESRRLTNGKYSYINKLKF- 545
Db 813 V-AQPPVVDVAAGGIDIASRKRGR-PALTFTTFLAAPGEBQRRRTKGRSEFTTFVEIV 870
Qy 546 GRVKQWQTGEASSKLDPSKVIORVAETEGT---DEIGLIYNKAGNDDIFVGGQKNID 603
Db 871 GKQDRWRIRDGAADTTIDLAKVWSQLVDANGVLKHSIKLDVIGGDDGVVLANASRIHYD 930
Qy 604 GGDGHDVFSKDGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKQETKVGKRTTE 663
Db 931 GGAGTNTVSVAALGRQSDITV-----SADGERFNRKQLNANVREGVATOTTAYGKRTTE 986
Qy 664 TIQYRDVELRVGYGYOSTDLNKLKSVVEVIGSQFNDVPKGSFNFIHSGEGDLDLGGAG 723
Db 987 NVQYRHVELARVGQVVE-VDTLEHVQHIIGGAGNDSITGNAHDNFLAGSGGDDRLDGGAG 1045
Qy 724 DDLRFGGKG----- 732
Db 1046 NDTLVGEGQNTVIGAGDDVFLQDLGVWSNQLDGGAGVTVTKYNNVHPQSEERLERMGDT 1105
Qy 733 -----NDRLSGDEGDDLLDGGSGDDVNLNG 756
Db 1106 GIHADLOKGTVEKWPALNLSFVDHVKNJENLHGSRLNDRAGDDQDDELWGHGNDTIRG 1165
Qy 757 GAGNDVIFRKGNDNTLYDGTGNDKLAFAFANISDIM-----IERTKEGIIVKRNDSHG 811
Db 1166 RGGDD--ILRGGGLGLTYLGEDGNDIFLQDDTETVSDDIDGGAGLTDVDYSAMI-----HPG 1219

Qy 812 SINIPRWY---ITSNLQYQSNKTD-----HKIEOLIG---KDGSVITSDDIDKI 855
Db 1220 RIVAPHEYGEGIEADLSREWRKASALGVDDYDNRVNVNIGTSMKD-VLIGDAQANTL 1278
Qy 856 LQDKKQDGTV 864
Db 1279 MGQGGDDTV 1287

RESULT 16.
CYAA BORBR ID_CYAA BORBR STANDARD; PRT; 1705 AA.
AC Q57506; O05179;
DT 01-NOV-1997 (Rel. 35, Created)
DT 18-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bifunctional hemolysin-adenylate cyclase precursor (Cyclolysin) (ACT)
DE (AC-HLY) [Contains: Calmodulin-sensitive adenylate cyclase
DE (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl cyclase); Hemolysin].
GN CYA OR CYAA.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CIP 9.73;
RX MEDLINE=96009899; PubMed=7557410;
RA Betson F., Simeiro O., Danchin A., Guiso N.;
RT "Cloning and sequence of the Bordetella bronchiseptica adenylate
RT cyclase-hemolysin-encoding gene: comparison with the Bordetella
RT pertussis gene.";
RL Gene 162:165-166 (1995).
RN [2]
RP REVISIONS TO 1555-1558.
RA Danchin A., Boursaux-Bude C.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS ADENYLATE CYCLASE BELONGS TO A SPECIAL CLASS OF
CC BACTERIAL TOXIN. IT CAUSES WHOOPING COUGH BY ACTING ON MAMMALIAN
CC CELLS BY ELEVATING CAMP-CONCENTRATION AND THUS DISRUPTS NORMAL
CC CELL FUNCTION.
CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC -1- ENZYME REGULATION: ACTIVATED BY HOST CALMODULIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
CC ACTIVITY (BY SIMILARITY).
CC -1- PTM: RELEASED IN A PROCESSED FORM.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE ADENYLYL
CC CYCLASE CLASS-2 FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE RTX
CC PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z37112; CAA85481.2; -.
CC HSP; P40136; IK90.
CC InterPro; IPR005165; Anthrax_toxa.
CC InterPro; IPR001343; Hemolysin_Ca_bind.
CC InterPro; IPR003995; RtxA.
CC Pfam; PF03497; Anthrax_toxa; 1.
CC Pfam; PF00353; hemolysinCbind; 17.
CC Pfam; PF02382; RTX; 1.
CC PRINTS; PR00313; CABNDNGRPT.
CC PRINTS; PR01488; RTXTOXINA.
CC PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 5.

Query Match 17.2%; Score 801.5; DB 1; Length 1705;
Best Local Similarity 27.7%; Pred. NO. 1.1e-32;
Matches 254; Conservative 135; Mismatches 336; Indels 193; Gaps 28;

QY 106 GLDSVENIDRKLKASNVLSLTSSFLGTALAGIELDSLI-----KKGDAAP----- 151
DB 403 GYDSDLGV-----GSRFSFLGEVSDMAVAEAELEMTQVHLHAGARQDDAEPGVSGASAHW 458

QY 152 -----DALAKASIDLINEIGNLSQ-----STQF-IEAFSSQLAKLG-----STI 190
DB 459 CORALOGAQAQVAARQLVHALMTQFGRAGSTNTQPEASLSRAAVFGLGASSAFAETV 518

QY 191 S-----QAKGFNIGNKQLNLFNSKTNIGLEIITGLSGISAGFALADKNASTGKKV 242
DB 519 SGFFRGSRRWAGGFGVAG-----GAMALGGGIGAVGAGMSLTD-DAPAGQKA 564

QY 243 AAGFELSNOVIGNVTKATSSVYL-----AQRVAAGLSTTGAVAAITSSIMLAISLAPMN 298
DB 565 AAGABIALQLTGTVELASSIALAARAVTSGLOVAGSAGAAAGAAALASPMETIYG 624

QY 299 AADKFNHANALDEFKQFKFGYGDHLLAEYQRCGVGTIEASLTTISTALGAVSAGVSA 358
DB 625 LVQQSHYADQLKLAQESSAVGYEGDALLAQLYRDKTAEGNAGVGSVLTVTGNAVSA 684

QY 359 AVGSVAGTPTIALVAGVTGLISGLEASKQAMFESVANRLOCKILEWEKONGQNYFDKG 418
DB 685 AAASVVGAPVAVVTSLLTALNGILRGVQPIIEKLANDYARKI---DELGGPQAYFEKN 741

Lyase; CAMP biosynthesis; ATP-binding; Hemolysis; Toxin; Virulence;
Whooping cough; Calcium-binding; Repeat; Lipoprotein; Palmitate.
CHAIN 1 312
CYCLASE.
CHAIN 313 1705
HEMOLYSIN (BY SIMILARITY TO E. COLI
HEMOLYSIN HLVA).
A, CATALYTIC.
B, ALA/GLY-RICH.
C.
D, ASP/GLY-RICH.
ATP (POTENTIAL).
28 X 5 AA REPEATS, GLY-RICH.
1.
2.
3.
4.
5.
6.
7.
8.
9.
10.
11.
12.
13.
14.
15.
16.
17.
18.
19.
20.
21.
22.
23.
24.
25.
26.
27.
28.
PALMITATE (BY SIMILARITY).
PALMITATE (BY SIMILARITY).
PALMITATE 1705 AA; C43B30F5886C835A CRC64;

QY 419 YDSRYAAYLANN---LKFLSELNKELEAERVIAITQORWNNNIGELAGITKLGRISGK 475
DB 742 LQARH-EQLANSGLRKLWADLQAGWNASSVIGVQTTEISKSALALEAAITGNADNLKAD 800

QY 476 AYADAFEDGKKVEAGSNITLDKTIIDISNSNGKTKQALHFTSPLTAGTESRRTNG 535
DB 801 VFVDRFIOGERV-AGQPWLDVAAGGIDIASRGER-PALTFTPLAAPGEEQRRRTKTG 858

QY 536 KYSYINKLKP-GRVKNQVVTGDEASSKLDFSKVIQORVAETEGT--DEIGLVNNAKAGND 592
DB 859 KSEFTTEVEIVGQDRWRIEDGAADTTIDLAKVVSQVLVDANGVLKHSIKLEIVGGDDV 918

QY 593 IFVGGKMNIDGGDGHDRVFYSKDGFGNITVDGTSATEAGSYTVNRKVARGDYIHEVVK 652
DB 919 VLNASRIHYDGGAGTNTVSYAALGRQDSITV---SADGERFNVKQLANNVYREGVA 974

QY 653 RQETKVGKRTETIQRYDYELRKVGYGYOSTNLKSVEEVIGSQPNDFVKSGKFNDFHSG 712
DB 975 TQKTAYGKRTENVQYRHVELARVQ-LVEVDTLHVQHIIIGGAGNDSITGNAHDFLAGG 1033

QY 713 EGGDLLGGAGDDRLFGCKGNDRLSGDEGDL-----LDGSG----- 750
DB 1034 AGDRLDGGAGNDTLVGEGHNTVVGAGDDVFLQDLGVMSNQLDGGAGVDTVKYNVHQ 1093

QY 751 -----DDVL 754
DB 1094 SEERLERMGDTGIHADLQKGTVEKWPALNLFSDVHVKNENLHGSSLNDSTAGDDRDEL 1153

QY 755 NGGAGNDVY-----IFRKGDGNDTLVDGTGNDKLAFAADANISDIM-----IERTKEGI 802
DB 1154 WGGDGNDTIHGRGGDDILRGGGLGLTLVDGNDIFLQDDTVSDDDIGGAGLTDVDSA 1213

QY 803 IVKNDHSGSINIPRWY---ITSNLQNYQSNKTDHK-----TEQLIG---KGSY 846
DB 1214 MI-----HAGKIVAPHEYFGIEADLSEGWKRAARGMDYDVSRSVENVIGTSMKO-VL 1268

QY 847 ITSQIDIKILQDKKQDGV 864
DB 1269 IGDAQNTLMGQGGDDTV 1286

RESULT 17
FRPC_NEIMB STANDARD; PRT; 1829 AA.
AC Q9JYU5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Iron-regulated protein frpC.
GN FRPC OR NM01415.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175753; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A., Mason T., Ciecko A., Parksey D.S., Blair E., Citron H., Clark E.B., Cotton M.D., Ufferback T.R., Khouri H., Qin H., Vamathevan J., Gill J., Scarlato V., Maignani V., Pizzo M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RA "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58";
RL Science 287:1809-1815 (2000).
CC -!- FUNCTION: May participate in the pathogenesis of meningococcal disease.
CC -!- SUBCELLULAR LOCATION: Outer-membrane associated and secreted (By similarity).

Db 1202 -----SDRLYAYQSGNTLNGSLGDDYLVG 1225

RESULT 18

FRPC_NEIMC
ID FRPC_NEIMC STANDARD; PRT; 1829 AA.
AC P55127;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Iron-regulated protein frpc.
GN FRPC.
OS Neisseria meningitidis (serogroup C).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=135720;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FAM20 / Serogroup C;
RX MEDLINE=94018616; PubMed=8412674;
RA Thompson S.A., Wang L.L., Sparling P.F.;
RT "Cloning and nucleotide sequence of frpc, a second gene from
RT Neisseria meningitidis encoding a protein similar to RTX
RT cytotoxins.";
RL Mol. Microbiol. 9:85-96(1993).
CC -!- FUNCTION: May participate in the pathogenesis of meningococcal
CC disease.
CC -!- SUBCELLULAR LOCATION: Outer-membrane associated and secreted.
CC -!- DOMAIN: The Gly-rich region is probably involved in binding
CC calcium, which is required for target cell-binding or cytolytic
CC activity.
CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L06299; AAA99902.1; -.
DR PIR; S35027; S35027.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtXA.
DR Pfam; PF00353; hemolysinCabin; 22.
DR PRINTS; PRO0313; CABNDNGRPT.
DR PRINTS; PRO1488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCIUM; 17.
KW Toxin; Calcium; Outer membrane; Repeat.
FT DOMAIN 879 1702 43 X REPEATS, GLY-RICH.
FT REPEAT 879 884 1.
FT REPEAT 888 893 2.
FT REPEAT 897 902 3.
FT REPEAT 1016 1021 4.
FT REPEAT 1025 1030 5.
FT REPEAT 1034 1039 6.
FT REPEAT 1043 1048 7.
FT REPEAT 1052 1057 8.
FT REPEAT 1061 1066 9.
FT REPEAT 1070 1075 10.
FT REPEAT 1079 1084 11.
FT REPEAT 1088 1093 12.
FT REPEAT 1097 1102 13.
FT REPEAT 1216 1221 14.
FT REPEAT 1225 1230 15.
FT REPEAT 1234 1239 16.
FT REPEAT 1243 1248 17.
FT REPEAT 1252 1257 18.
FT REPEAT 1261 1266 19.
FT REPEAT 1270 1275 20.
FT REPEAT 1279 1284 21.

FT REPEAT 1288 1293 22.
FT REPEAT 1297 1302 23.
FT REPEAT 1416 1421 24.
FT REPEAT 1425 1430 25.
FT REPEAT 1434 1439 26.
FT REPEAT 1443 1448 27.
FT REPEAT 1452 1457 28.
FT REPEAT 1461 1466 29.
FT REPEAT 1470 1475 30.
FT REPEAT 1479 1484 31.
FT REPEAT 1488 1493 32.
FT REPEAT 1497 1502 33.
FT REPEAT 1616 1621 34.
FT REPEAT 1625 1630 35.
FT REPEAT 1634 1639 36.
FT REPEAT 1643 1648 37.
FT REPEAT 1652 1657 38.
FT REPEAT 1661 1666 39.
FT REPEAT 1670 1675 40.
FT REPEAT 1679 1684 41.
FT REPEAT 1688 1693 42.
FT REPEAT 1697 1702 43.
SQ SEQUENCE 1829 AA; 197622 MW; 5C3494020A782DC8 CRC64;

Query Match 9.4%; Score 435; DB 1; Length 1829;
Best Local Similarity 21.9%; Pred. No. 3.2e-14;
Matches 231; Conservative 167; Mismatches 355; Indels 302; Gaps 48;

Qy 43 DFIKAA-----ELGIARLAEPEPHETAKSVDTVNQFL-----SLTQTGLAISATKLEKF 94
Db 478 DNTKLADGSPAKHGAYAAELDSNGDNIINAADAAAFQTLRWQDLNQDGIS----- 528
Qy 95 LQKHSNKLAK-GLDSE-----NIDRKLKASNVLSLTSSFLGTALAGIELDSLKKGD- 148
Db 529 -QANELRTLEELGQSLDLAYKDVKNLGN-GNTLAQOGSYTKDTGTAKMGDLALLADN 586
Qy 149 -----AAPDALAK--ASIDLINEI-----IGNLSQSTQTIIEAFSSQLAKL 186
Db 587 LHSFKDKVELTAQAKAANLAGIGRLDLREAAALSGDLANMLKAYSAATKEAQLALL 646
Qy 187 GSTISQ-AKGSFNIGNKQ---NLNFSKT-NLGLIITGLLSGI--SAGPALAK----- 234
Db 647 DNLIHKWAETDSNMGKSPMRLSTDWTQTANEGIALTPSOVAQLKQNALVSLSDKAKAAI 706
Qy 235 -----NASTGKKVAAGFELSNOVIGNVTKAISSYV--LAORVAAGLSITTCGAVAL 282
Db 707 DAARDRIAVLDAYTGQDSSTLYNSEEEDALNIVKVTNDTYDHLAKNIYNL-----L 758
Qy 283 ITSSIMLAISPLAFMNAAD-----KFNHA-----NALDEFKQFR-----KFGY 321
Db 759 FQTLQPYLNQISPKMENDITFLDFSGLVQAFNHVKTNPQKAFVDLAEMLAYGELRSWY 818
Qy 322 DGDHLLAEY---ORGVTIE---ASLTTISTALGAVSAGVSAAGVAVGPIALLVA-- 373
Db 819 EGRELMADYVEEAKKAGKFEDYQVLQGETVALLAKTSGTOADDILQNVGFGHNKNSLY 878
Qy 374 ---CVTGLISGILEASKOAMFESVANRLQKLEWEKQNGQNY-FDKG-----YDSRY 423
Db 879 GNDGNDTLIGG---AGNDYL-----EGSGSDTYVFGFGQDVTYNYDY 920
Qy 424 AAYLANNLKLSELNKL-----EAERVI-----AITQORWNNIG----- 459
Db 921 ATGRKDIIRFTDGITADMLATFTREGNHLIIKAKDDSGQVTQVSFYQNDGSGAYRIDEIHF 980
Qy 460 -----ELAGITKLGRIKSGKAYADAFEDGKKVEAG-----SNITLDAKTGLIID 503
Db 981 DNGKVLDAVATVKELVQOSTDGSRLYAYQSGNTLNGGLGDDYLYGADGDDLLNLDAGNDS 1040
Qy 504 ISNSNGKKTQALHFTSPLLTAGTESRRLTNGKYSYINKLKFGVRKQVQVTDGASSKLD 563
Db 1041 IYSGNGNDT-----LNGGEGNDAL-----YGNNGDALNGG----- 1071
Qy 564 FSKVIQORVAETEGTDEIGLIVNAKAGNDIDFVGQGWNIIDGGDGHDRVYFSKDGFGNIT 623

Db 1072 -----EGNDHL-----NGEDGNDTLIGGAGNDYLEGSGSDTYVFGK--GFQDQT 1114
Qy 624 V-----DGTSA-----TEAGSY--TVNRKVARGDYIHEVVKRQETKVGKR 661
Db 1115 VYNYDYATGRKDIIRFTDGTADMLTFTREGNHLIIKAKDGGSGQVTVQSYFQNDGSGAYR 1174
Qy 662 TETIOYRDYELRVG-----YGVQSTDNLK---SVEEVIGSOFNDVFKG 702
Db 1175 IDEIHFNGKVLVATVKELVQOSTDGSRLYAYQSGNTLNGGLGDDYLVGAGDDLLNG 1234
Qy 703 SKFNDFHSGEGDLDLPGAGDDRFL-----GGKGNDRLSGDEGDDLLDGGSDDV 753
Db 1235 DAGNDSIYSGNDTDLGGEGNDALYGVNGDALNGGEGNDHNGEDGNDTLIGGAGNDY 1294
Qy 754 LGGAGNDVYIFRKGNDTLYD---CTG-NDKLAFADANISIMIERTEG--IIVKRN 807
Db 1295 LEGSGSDTYVFGKFGQDVTYNYDYATGRKDIIRFTDGTAD-MLTFTREGNHLIIKAK 1353
Qy 808 DHSGSINIPRWYITSNLQYQSNKTDHKIEQLGKGSYITSDOIKILODKKDGTVITS 867
Db 1354 DDCQVTVQSYF-----QNGSGA--YRIDEIHFNGKVLVATVKELVQOSTDG----- 1401
Qy 868 QELKKLADENKQSKLASDIASSINKLVGSMALFG 902
Db 1402 -----SDRLYAYQSGSTLNGGLGDDYLVG 1425

RESULT 19

FRPA_NEIMB STANDARD; PRT; 1302 AA.
ID FRPA_NEIMB STANDARD; PRT; 1302 AA.
AC OSK0K9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Iron-regulated protein frpa.
GN FRPA OR NMB0585.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58";
RL Science 287:1809-1815 (2000).
CC -!- FUNCTION: May participate in the pathogenesis of meningococcal
CC disease.
CC -!- SUBCELLULAR LOCATION: Outer-membrane associated and secreted (By
CC similarity).
CC -!- DOMAIN: The gly-rich region is probably involved in binding
CC calcium, which is required for target cell-binding or cytolytic
CC activity.
CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.ebi.ac.uk/ebis-sib.ch/](http://www.ebi.ac.uk/ebis-sib.ch/announce/)
CC or send an email to license@sib.ch).
CC -----

DR EMBL; AE002414; AAF41013.1; --
DR PIR; C81182; C81182.
DR TIGR; NMB0585; --
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 12.
DR PRINTS; PR00313; CABINDNGRPT.
DR PRINTS; PR01488; RXTXOXINA.
DR PROSITE; PS00330; HEMOLYSIN CALCIUM; 9.
KW Toxin; Calcium; Outer membrane; Repeat; Complete proteome.
FT DOMAIN 752 975
FT REPEAT 761 766 2.
FT REPEAT 770 775 3.
FT REPEAT 789 794 4.
FT REPEAT 803 808 5.
FT REPEAT 812 817 6.
FT REPEAT 821 826 7.
FT REPEAT 830 835 8.
FT REPEAT 843 848 9.
FT REPEAT 852 857 10.
FT REPEAT 861 866 11.
FT REPEAT 870 875 12.
FT REPEAT 879 884 13.
SQ SEQUENCE 1302 AA; 141397 MW; 21D058C5C98BDE8 CRC64;
Query Match 9.1%; Score 421; DB 1; Length 1302;
Best Local Similarity 22.5%; Pred. No. 1e-13;
Matches 198; Conservative 136; Mismatches 292; Indels 254; Gaps 39;
Qy 193 AKGFSNIGNKLNLFNFKTNLGLIITGLSGISAGFALADKNASTGKKVAAGFELSNQV 252
Db 303 AKGFS-----GSLFDHTNNGIRTATGWYSA-DDGLLVRLDN-----GNGI 341
Qy 253 IGNTVKAISYVLAQRVAAGLSTTGVAALITSSIMLAISPLAFMNAADKENANALDEF 312
Db 342 IDNGAELFGDNT---KLADGSFAXHGYAALAEID-----SNGDNIINGAADAFAQS 388
Qy 313 AKOFKRFYDGD---DHLLAEYQGVGTTEASITTTISTALGAVSAGVSAAGVAVGTPI 368
Db 389 LRWQDLNQDGIQANELRTLELGIQSLDLAYKDVNKNLG---NGNTLAQOQSYTKTD- 444
Qy 369 ALLVAGVTGLISGLEASKQAMFESVANRLQKILEWEKQNGQNVFDKG--YDSRYAAY 426
Db 445 -----GTTAKMGDLLLLAA-----DNLHRSFKDKVELTAEQAANLAGIGRLRLREAAA 494
Qy 427 LANNLFELSELNKELEA-ERVIAITQ---QRW---DNNIGELA-----GIT 465
Db 495 LSGDLANMLKAYSAAEETKEAQLALLDNLIHKWAETDSNWKKSPPMLRSTDTWTOTANEGIA 554
Qy 466 KLGERI-----KSGKAYADAFEDGKKVEAGSNITLDAKTG----- 500
Db 555 LTPSQVQALKKNALVSLSDKAKAIAIDARDRIAV-----LDAYTQDSNTLYMSEED 607
Qy 501 ---IIDISNSN-----GKKTQALHFTSPLLTAGTES 528
Db 608 ALNIVKVTNDYDHLAKNIYQNLFFQRLQPYLNQISFKMENDFTFLDF-SGLVQAFNHV 666
Qy 529 RERLTNGKYIYN---KLKFRGVKNW-----QVTDGEASSKL-DFSKVIR- 570
Db 667 KE--TPQKAFVDLAEMLAYGELRSWYEGRRRLMTDYVEEAKKAGKFKEDYQKVLGOETVAL 724
Qy 571 VAETEGTDEIGLIVN-----AKAGNDIFFVGGCKMNIIDGGDGHDRVRFYKDDG 618
Db 725 LAKTSQTQADDILQNVGFHKNKVSLYGNDGNDTLGGAGNDYLEGSGSDTYVFG--G 782
Qy 619 FGNITV-----DGTSA-----TEAGSY--TVNRKVARGDYIHEVVKRQET 656
Db 783 FGQDVTYNYDYATGRKDIIRFTDGTADMLTFTREGNHLIIKAKDGGSGQVTVQSYFQNDG 842
Qy 657 KVKRRTETIQYRDYELRVG-----YGVQSTDNLK---SVEEVIGSOFNDVFKG 697
Db 843 SGAYRIDEIHFNGKVLVATVKELVQOSTDGSRLYAYQSGNTLNGGLGDDYLVGAGDD 902

DE Reticulocyte binding protein 1 precursor.
GN RBPI.
OS Plasmodium vivax (strain Belem).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN SEQUENCE FROM N.A.
RP MEDLINE=92315338; PubMed=1617731;
RX Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RA "A reticulocyte-binding protein complex of Plasmodium vivax
RT merozoites";
RT Cell 63:1213-1226(1992)
RL CC -!- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC HUMAN RETICULOCYTE CELLS.
CC -!- SUBUNIT: Homodimer (potential).
CC -!- SUBCELLULAR LOCATION: Membrane-bound.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M88097; AAA29743.1; -
KW Malaria; Receptor; Signal; Transmembrane.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 2869 RETICULOCYTE BINDING PROTEIN 1.
FT DOMAIN 18 2807 EXTRACELLULAR.
FT TRANSMEM 2808 2826 POTENTIAL.
FT DOMAIN 2827 2869 CYTOPLASMIC.
FT SITE 1030 1032 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 2599 2601 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 2869 AA; 330213 MW; B9DBE442205EBCFF CRC64;

Query Match 4.1%; Score 188.5; DB 1; Length 2869;
Best Local Similarity 19.7%; Pred. No. 0.14;
Matches 209; Conservative 175; Mismatches 360; Indels 315; Gaps 50;

QY 2 SNINVIKNIQA-----GLNSTKGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGI 53
DB 1214 STANTLKNAGKENEHLELNKTKGQMDIYEKLIABELKEGTWNE-LKDA NEKA- 1271

QY 54 ARLAEPNHT-----ETAKKSVPTVNOFLSLTGTGTAISATKLEFLQKHS- 99
DB 1272 NKVEPEPERNIIGHVLRTITVEKDAGKGVVEENSL-----KTKIEKLIQETSD 1320

QY 100 -----TNKLAKGLDSV---ENIDRKLKASNVLSLSSFLGTALAGIELDSLKKGDA 149
DB 1321 DSQNELVTTSTIKHLENAGYEDVIKRNEDSIQREKAKSLET-----LDEMCK---- 1370

QY 150 APDALAKASIDLINRIIGNLSQSTOTIEAFSSOLAKLGSSTISQ--AKGFSNI-----GNK 202
DB 1371 -----LVQVNNQLSAIQGNAGISKELNKGVIELLIISTYNSILEVVKNS 1419

QY 203 LQNLNFSKTNLGLLEITGLSGISAGFALADKNASTGKVAAGFELSNOVIGNVTKAIS 262
DB 1420 SESVRFSQ-----LANGEFTKAEGBEKNASA--RLAEAKLKEQIV-----KLDY 1463

QY 263 VYLAQVRAAGLSTTCGAVALITSSIMLAISPLAFNVAADKF-----NHANALDEPAK--Q 315
DB 1464 SDIDDKVK-----KIEGKREILKMKESALTTFWEESEKFKQMCSSHMENAKGKKKIE 1516

QY 316 FRKFGYDG-----DHLAEYQGVGTTEASLTITSTALGAVSAGVSAAGVAVGTPIA 369
DB 1517 YLKNNGDGGKANITDSQMEVGNVYKAEHAFHTVEAQVDTKAPCES----- 1564

QY 370 LLVAGVTGLISGILBASKOAMPESVANRLOQKILEWKONGQNVFD---KGYDSRYAAY 426
DB 1565 -IVAVYTKV-----DNLFNLSLMKEVKVC-----EKKNDKAEKYSAKLPYDGRIKAR 1612

QY 427 LANNLKLSELNKELEAER-----VIAITQQRWD---NNIG----- 459

Db 1613 VSENERKISELKAKVKEKSSOLNDVSTKSLQIDNCRQOLDVLSNIGRVKQNALQY 1672
QY 460 -----ELAGITKIG-----ERIKSGK-AVADAFEDGK-----KVZAGSNITLDK 498
DB 1673 FDSADKSMKSVLPISLGAESKSLDKVKAAKESYKKNLETQVNMESRINVEEGSLTDIDKK 1732
QY 499 TGIIDISNSNGKTKQALHFTSPILTAGTESRRLTNGKYSYINKLKEG---RVKNQWQVTD 555
DB 1733 --ITDIEND-----LLKMKKQYEEGL-----LQIKENADKRSKNFELVG 1770
QY 556 GEASSKLDKSVKIQORVAETEGTDEIGLIVNAKAGNDIFFVGGQKKNITDGGDGHDRVFSYK 615
DB 1771 SEINALDLPSTSFIFKLKKEYDWTGLKNYGV-----KQNEIHGE-----FTK 1814
QY 616 DGGFNGITVDGTSATEAGSYTVNRKVARGDYIHEVVKRQETKVGKRTTETIQRYDYELRKV 675
DB 1815 -----SYNLIETHLSNATDYSVTPEKAQS--LRELAKEBEHLRRREEAIFLLNDIKV 1867
QY 676 GYGQSTDNKSVSEEVIGSQFNDVPKSGKNDIFHSGEGLDLDGGAGDDRLFGGKGNDR 735
DB 1868 -----ESLKLKEMM-KKVSAYEGMKRDHTSVSOLVQDM-----K 1902
QY 736 LSGDEGDDLDGGSGDDVNLGGAGNDVYIFRK-----GQNDTLYDGTGNDKLAIF 785
DB 1903 TIIVDELKTLNDISECSSVLN-----NVSVIKVKVSKHADYRRDAN-SMYE---SMVTL 1953
QY 786 ADANISD-IMERTKEGIIKVRNDHSGSINIPRWYITSNLQYNSKNTDHHKIQ----- 838
DB 1954 ANYFLSDEAKITSSGMFEPAEMKSNFKTDLEIFSVISN-----SNELLKKIEQSDNVI 2008
QY 839 -----LIGKGSYI-----TSQIDKILKODKOGTIVTSQE-----LKKLAD----- 875
DB 2009 QKERESQLAKADTDIYVNIKLVKNEFKLEAKNKEEVSEKVRKALRKLRSQVEGIRCH 2068
QY 876 -ENKSQKLSASDIASSLNKLGVSMALFGTANSVSSNALQ 913
DB 2069 FENFRLLDNTELENLKKWV---TIYRDKKSERESGLQ 2104

RESULT 23
HLVA PROMI STANDARD; PRT; 1577 AA.
AC P16466;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE Hemolysin precursor.
GN HPMA.
OS Proteus mirabilis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Proteus.
OX NCBI_TaxID=584;
RN
RP SEQUENCE FROM N.A. AND SEQUENCE OF 30-43.
RC STRAIN=Isolate 477-12;
RX MEDLINE=90170827; PubMed=2407716;
RA Uphoff T.S., Welch R.A.;
RT "Nucleotide sequencing of the Proteus mirabilis calcium-independent
RT hemolysin genes (hpmA and hpmB) reveals sequence similarity with the
RT Serratia marcescens hemolysin genes (shlA and shlB).";
RL J. Bacteriol. 172:1206-1216(1990).
CC -!- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC DEFINED.
CC -!- FUNCTION: CELL-BOUND HEMOLYSIN, WHICH RELEASES HEME-IRON FROM
CC ERYTHROCYTES BY INTERACTION WITH THE ERYTHROCYTE MEMBRANE. HPMA
CC REQUIRES HPMB FUNCTION.
CC -!- SUBCELLULAR LOCATION: Outer membrane.
CC -!- MISCELLANEOUS: THE CONSERVED AMPHIPATHIC DOMAINS IN SHLA AND HPMA
CC MAY BE RESPONSIBLE FOR PORE FORMATION.
CC -!- SIMILARITY: TO S.MARCESCENS HEMOLYSIN (SHLA).
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M30186; AAA25657.1; -;
DR PIR; A35140; Outer membrane; Signal.
KW Hemolysis; Toxin; 29
FT SIGNAL 1
FT CHAIN 30 1577 HEMOLYSIN.
SQ SEQUENCE 1577 AA; 165869 MW; 175975E0C924B2D9 CRC64;

Query Match 3.9%; Score 183.5; DB 1; Length 1577;
Best Local Similarity 20.3%; Pred. No. 0.11;
Matches 211; Conservative 132; Mismatches 371; Indels 325; Gaps 53;

QY 3 NINVIKSN-----IQAGLNSTKSGKLNLYLAIPKDYDPQKGLTNDLFKAA 48
DB 509 NLNVQKTNNDKVTVDNHHVMGGIGGQGNKNNNQVSHATQLTAD----- 554
QY 49 DELGIARLAEPEPHNTAKSVDTVNOFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLD 108
DB 555 ---GQLLADNNVNIIGTSQVKGNGQAFVKTQGDVID-----NALS---E 595
QY 109 SVENIDRLKASNVLTSLGFLGTALAGIELSLIKKGAAPDALAKASIDLIN-----E 164
DB 596 TISKIDERTGAFNI--TKSS-----HKNETNKQTSGLSLISDAQLTVVSGNDVN 644
QY 165 IIGNLSOSTQTIAPFASQLKGLSTISQAKFSNI--GNKLQNLNFKSTNGLLEIITGLL 222
DB 645 VIGSLINKSAD-----KLG---IHSGLDINVKSAQQVTKIDDEKTSLSA---ITGHA 688
QY 223 SGI-----SAGFALADKNASTGKVAAGFELSNOVIGNVTKAISSVYLAORVAAGLSTTG 277
DB 689 KEVEDKOYSAGFHITH--FTNKYTSITETEQA NSTISGANVDLOANKVDVTFAGSDAKTTA 745
QY 278 AVAALITSSIMLATSPLAFMNAADK-----FNHANLDBFAKQFRKFGYDGD 324
DB 746 GNASITGDNV-----AFVSTENKKQTDNTDTTISGFSYTGVDKVGSK-ADFQYDKQ 797
QY 325 HLLAEY--QRGVGTIEASLTITISALGNVSAAGVAAVG-----SAVGT 366
DB 798 HTQTEVTKNRSQTEVAGDLTITANKOLLHEGASHVEGRYQESGENIQHLAVNDSETSK 857
QY 367 PIALLV---AGVTGLISGILEASKOAMPESVANRLOKLEWEKQNGQNYFDKGYDSRY 423
DB 858 TDSLNVGIDVGNLDYSGVTKPVKKAIEDGVNTTKPG----- 894
QY 424 AAYLANNLKFLSELNKELEAERVIAITQORWDDNNIGELAGI-----TKLGERITKSKAY 477
DB 895 -----NN---TDLTKKVTARDAIA-----NLNLSNLETENVGVEVG--IKGGS- 934
QY 478 ADAPEDKKVPAGSNITLDATGTLIDISNSNGKKTQALHTSP-----LLTAGTESRELTL 533
DB 935 QQSQTDSQAVSTSN-----AGKIDIDSNKLNKHDQGYTHYQSTQEGISLTANTHTSEATL 988
QY 534 NGKYSYINLKFGRVGNQWQVTDGE--ASSKLDKFKVIQORVAETEGTDEIGLIVNAK---- 587
DB 989 DXQHTTHETKGG-----GQIGVSTKTKGSDITVAIKEGGQTTDNALMETKAKSQF 1039
QY 588 AGNDIDIFVG-----QKGNIDGG-----DGHDRVFFVSKDGGFGNIT 623
DB 1040 TSNGDISINVENGAHYEQAQDAQKGTVINAGGDLTLAQATDTTSE--SQSNVNGSANLK 1098
QY 624 VDGTL--SATEAGSVTVNRKARGDIYHEVVKQST-KVKKKTETIQYRDYELRKVGVGQ 680
DB 1099 VGTTPESKDYGGGFNA-----GTTTHS--KEQTTAKVGTITGS---QGTELNA---GHN 1144
QY 681 ST---DNLSKEEVIQSGFNDV-----FKGSKFNDIFHSGEGDGLLD-----GGA 722

DB 1145 LTLOGLTHLSSEQDIALNATNKVDLQASSEHTEKGNLGGVQAGFGKKMTDDASSVNGL 1204
QY 723 GDRRLFGGKGNRLSDGEGDGLLDGSGDDVLNGAGAGNDVYIFRKGNDNTLDYDGTGNDK 782
DB 1205 GSAQFAIGKDEKSVSREGGTI--NNSGNLTIN---GNSVH----- 1240
QY 783 LAPADANISDIMEIERTKEGIIVKRNHDSGSIINIPRWYITSNLQYQSN-KTDHKIEQLIG 841
DB 1241 LQGAQVNSKDTQL-----TSQSGDIEI-----TSAQSTDYKKNWCTD-----IG 1279
QY 842 KQGSYITSDQIDKILQDKKQDGTVTISQELKGLADENKSK-----LSASDTASSLNKL 894
DB 1280 FNKG--KTNTPTKEVTEEPATSIHNIGGKLLVNVVEDQOKTSHQATLETGTLTINSNK- 1336
QY 895 VGSMALEFG---TANSVSSN 910
DB 1337 --DLTSLGANVTADSVTGN 1353
RESULT 24
YL24_ANASP
ID YL24_ANASP STANDARD; PRT; 1693 AA.
AC Q8YV57;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical WD-repeat protein all2124.
GN ALL2124.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
CC -|- SIMILARITY: Contains 13 WD repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP003588; BAB73823.1; -;
DR PIR; AF2071; AF2071.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 14.
DR PRINTS; PR00320; GPROTEINBRPT.
DR SMART; SM00320; WD40; 14.
DR PROSITE; PS00678; WD REPEATS 1; 3.
DR PROSITE; PS00082; WD REPEATS 2; 14.
DR PROSITE; PS0294; WD REPEATS REGION; 1.
KW Hypothetical protein; Repeat; WD repeat; Complete proteome.
FT REPEAT 1070 1109 WD 1.
FT REPEAT 1111 1150 WD 2.
FT REPEAT 1153 1192 WD 3.
FT REPEAT 1195 1236 WD 4.
FT REPEAT 1280 1318 WD 5.
FT REPEAT 1320 1359 WD 6.
FT REPEAT 1361 1400 WD 7.
FT REPEAT 1402 1441 WD 8.
FT REPEAT 1444 1483 WD 9.
FT REPEAT 1486 1525 WD 10.
FT REPEAT 1528 1567 WD 11.


```
FT REPEAT 1570 1609 WD 12.
FT REPEAT 1611 1650 WD 13.
SQ SEQUENCE 1683 AA; 185459 MW; F7CB361FF54F7137 CRC64;

Query Match 3.9%; Score 180.5; DB 1; Length 1683;
Best Local Similarity 19.6%; Pred. No. 0.17;
Matches 197; Conservative 125; Mismatches 354; Indels 327; Gaps 44;

QY 31 KDYPQKGTLLNDFKAADELGIALAE--PNHETAKSVDTVNQPLSLTQTGIALISA 88
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 767 QSYQBLGG-----IKGALERSQGVYSLPQLOCAKWI-----FLSLTQLGEGTED 815
QY .89 TKLEFLQKHSNKLAKGLDSVENIDRKLGRASNVLSTLSSFLGTALAGIELDSLKKGD 148
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 816 TRRIYKSDLVKKYPAGL-----VEQTLNLTNAKLVVINLEAEI----- 856
QY 149 AAPDALAKASIDLINIEIGNLSQSQTTEAFSSQLAKLGSTI-----SQAKGFSNIG 200
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 857 ---EAQGS-----FSPTFPNLSPTPVTVEVAHEILIRHWSLRLWLEBDRRLRKQIN 909
QY 201 NKLQNLNFKTNLGLIEITGLLSISAGFALADKNASTGKKAAGFEL-----SNQVIGNV 256
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 910 HACQ-----LWQSGKQADFL-----QCARLAERAEIYIYWTDELGADV 949
QY 257 TKAISSYVL-----AORVAAGLSTTGAAALITSSIMLAISPLAFMNRAD 301
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 950 QEFITGACLAERKHQOLQAKNRKQRAVVALSVLGIA-----VSFGGLAYWQGRE 1001
QY 302 -KFHNANALDEFKFRKFGVGDHLLAEYQGVGTIEASLTITSTALGAVSAGVSA-A 359
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1002 AQFREIALNLSQA-----NLLSHQ-----LAALIASLKAAQ 1036
QY 360 VGSAGVTPIALLVAGVTGLISGILASQANFPE-SVANRLOK-----ILEWKQNGQNYF 415
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1037 VNHVIAVNNLKLATVTL-----QOALFEMQERNLECHKDGVISISIRDGOTIA 1088
QY 416 DKGYDSRYAAYLANNLKFLSELNKELEAEVIAI-----TORWNNIGEL 461
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1089 SGSLDKTKIKMSRDRGLFRT-LNGHEDAVYSVSPDQGTIASGSDKTIKLMQTSQTL 1147
QY 462 AGITKLGRIKSGKAYADAFEDGKKEVAGSN-----ITLDA-KTGIDIS 505
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1148 LK-TITGHEQTVNNVYFS--PDGKNLASDHSIKLMDTTSQQLMTLTGHSAGVTVR 1204
QY 506 NSNGKKTQALHFTSPLLTAGTESRE-----RLTNKYSYINKLKFG----- 546
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1205 FSPDQGT-----IAAGSEDKTVKLWHRQDGKLLKTLNGHQDWNLSLFSFGDKTLA 1255
QY 547 -----RVKNQVTDGEASSKL-----DESKVIQVRAETEGTDEI----- 580
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1256 SASADKTIKMWRIADGKLVTKLGNHDSVWNVFSSDGGKATASASRDNTIKLMNRHGIEL 1315
QY 581 -----GLIVNAK-----AGNDDI-----FVGQG 598
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1316 EFTTGHSGGVAVNPLPDSNIIASASLNTIRLMQRPILISPLEVLAGNSGVYAVSFLHDG 1375
QY 599 K-MNTDGGGDRVPSYKDGFP-----GNITVDGTSATEAGSYTVNRKVARGDYIHEVVK 652
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1376 SIATAGADGNLQIWLHSDGSLKTLPGNKAIYGISFTPGQDLIAS-----ANAD---KTVK 1429
QY 653 RQETKVGKRTETIQYRDYELRKVGY-----GYOSTNLKSVERVIGSQNDVFKGSKF 705
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1430 IWRVDRGKALKTLIGHDNEVKNVNFSPDGKTLASASRNTVKNVNSDGKFKTLKGTHT- 1488
QY 706 NDIHFGSGEDDLLDGGAGDRFLFGKGNDRLSGDEGDDLLDGGSGDDVLNAGAGNDVYIF 765
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1489 DEVFWVSPD-----GKIITASADKTIKRLWDSFGSLNKLKSLPAHNDLVYS 1535
QY 766 RKGDNLDLYDGTGNDKLAFAADANTSDIMERTKEGIIVKGNDSHGSNIPRWYITSLNQ 825
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1536 VNFNPDGSMLASTSADK-----TVKLWRSHDGHLL--HTFSGHSNVV----- 1575
QY 826 NYQSNKTDHKBQLGKDGYSITSDIKILQD-KKDGTVITS 867
```

Db 1576 -YSSS-----FSPDGRYIASASEDKTVKIWDIGHLLTT 1608

RESULT 25

NODO RHILV

ID NODO RHILV STANDARD; PRT; 284 AA.

AC P15728;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 01-MAR-1992 (Rel. 21, Last annotation update)

DE Nodulation protein O.

GN NODO OR NOLR.

OS Rhizobium leguminosarum (biovar viciae).

OG Plasmid sym PRLIJI.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.

OX NCBI_TaxID=387;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-8..

RC STRAIN=8401;

RX MEDLINE=901151607; PubMed=2303029;

RA Economou A., Hamilton W.D.O., Johnston A.W.B., Downie J.A.;

RT "The Rhizobium nodulation gene noda encodes a Ca2(+)-binding protein that is exported without N-terminal cleavage and is homologous to haemolysin and related proteins.";

RT J. Bacteriol. 171:6764-6770(1989).

RL EMO J. 9:349-354(1990).

RN [2]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 4-18.

RX MEDLINE=90078125; PubMed=2687250;

RA de Maagd R.A., Wijfjes A.H.M., Spaink H.P., Ruiz-Sainz J.E.,

RA Wijffelman C.A., Okker R.J.H., Lugtenberg B.J.J.;

RT "noda", a new nod gene of the Rhizobium leguminosarum biovar viciae sym plasmid PRLIJI, encodes a secreted protein.";

RL J. Bacteriol. 171:6764-6770(1989).

CC -1- FUNCTION: THE NODO PROTEIN MAY PLAY A ROLE IN NODULE DEVELOPMENT BY DIRECT INTERACTION WITH THE ROOT HAIR CELLS OR SOME OTHER PLANT SURFACE IN A CA(2+)-DEPENDENT MANNER.

CC -1- SUBCELLULAR LOCATION: SECRETED. BY A MECHANISM THAT DOES NOT INVOLVE AN N-TERMINAL SIGNAL PEPTIDE.

CC -1- INDUCTION: BY NARINGENIN (FLAVONOID).

CC -1- SIMILARITY: TO HEMOLYSIN AND RELATED PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X17285; CAA35178.1; -.

DR EMBL; M29532; AAA26341.1; -.

DR PIR; A43721; A43721.

DR PIR; S08385; S08385.

DR InterPro; IPR001343; Hemlysn_Ca_bind.

DR Pfam; PF00353; hemolysincabind; 5.

DR PRINTS; PR00313; CABNDNGRPT.

DR PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.

KW Plasmid; Nodulation; Calcium-binding; Repeat.

FT DOMAIN 2 73 8 X APPROXIMATE TANDEM REPEATS.

FT DOMAIN 92 127 4 X APPROXIMATE TANDEM REPEATS.

FT CA BIND 92 131 POTENTIAL.

FT DOMAIN 208 222 EXPORT SIGNAL (ASPARTIC ACID BOX (POTENTIAL)).

SQ SEQUENCE 284 AA; 30002 MW; 9CA41DCFCBDF8E15 CRC64;

Query Match 3.8%; Score 178.5; DB 1; Length 284;

Best Local Similarity 23.2%; Pred. No. 0.021;

Matches 61; Conservative 38; Mismatches 99; Indels 65; Gaps 9;

QY 684 NLKSVEE---VIGSQFNDVFKSGKFNDFHSGEGDILL---DG-----GAGDRLFGKG 731

Db 2 NIKSDNGSFKGSPENDIIDGGKNDWIDAGNGDDRIKAGDQDSITAGPGHDIVWAGK 61

Qy 732 GNDRLSGDEGDDLL-----DGGSGDDVINGAGNDVYIFKGDGNDT 773

Db 62 GSDVIHAGDGDLLYSASYLYVTDPRHVPISGEGDDVLYAGPGSDILV--AGDGDV 119

Qy 774 LYDGTGNDKLAFAADANISDIEMIET-----KEGLIVKRNHSGSINIPRWYIT 821

Db 120 LTGGDDGDAFV-----RFHDPWGTTHCYTSMVDPDTKQDFVLDADFGGRNL-----FD 172

Qy 822 SNLQNYQSN-----KTDHKIEQLGKGSYITSDQIDKILQKDGTVITS 867

Db 173 ANFINHSKGPGEFVDTFYNGAEGHGVVITDRGFASAAAAATAIDHEARGDIIVF 232

Qy 868 QELKLADEKSKLSASDIASS 890

Db 233 HDQKTLGQDGETHATLAYVDSA 255

RESULT 26

SLAP_CAUCR STANDARD; PRT; 1025 AA.

AC P35828; Q46015; Q9RF12;

DT 01-JUN-1994 (Rel. 29, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE S-layer protein (Paracrystalline surface layer protein).

GN RSAA OR CC1007.

OS Caulobacter crescentus.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;

OC Caulobacteraceae; Caulobacter.

OX NCBI_TaxID=155892;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-21; 646-660 AND 1021-1025.

RC STRAIN=ATCC 19089 / CB15;

RX MEDLINE=93007489; PubMed=1393820;

RA Gilchrist A., Fisher J.A., Smit J.K.;

RT "Nucleotide sequence analysis of the gene encoding the Caulobacter crescentus paracrystalline surface layer protein.";

RL Can. J. Microbiol. 38:193-202(1992).

RN [2]

RP REVISIONS TO 376; 636 AND 842-843.

RA Awram P.;

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=J33001;

RX Bingle W.H., Awram P.A., Nomellini J.F., Smit J.K.;

RA "The secretion signal of C. crescentus S-layer protein is located in the C-terminal 82 amino acids of the molecule.";

RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 19089 / CB15;

RX MEDLINE=21173698; PubMed=11259647;

RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock K.E., Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B., DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H., Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K., Uterback T., Tran K., Wolf A., Vamathevan J., Emolaeva M., White O., Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;

RT "Complete genome sequence of Caulobacter crescentus.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).

RN [5]

RP SEQUENCE OF 1-313 FROM N.A., AND SEQUENCE OF 1-20.

RC STRAIN=ATCC 19089 / CB15;

RX MEDLINE=89008089; PubMed=3049545;

RA Fisher J.A., Smit J.K., Agabian N.;

RT "Transcriptional analysis of the major surface array gene of Caulobacter crescentus.";

RL J. Bacteriol. 170:4706-4713(1988).

RN [6]

RP CHARACTERIZATION.

RC STRAIN=ATCC 19089 / CB15;

RX MEDLINE=98292737; PubMed=9620954;

RA Awram P., Smit J.K.;

RT "The Caulobacter crescentus paracrystalline S-layer protein is secreted by an ABC transporter (type I) secretion apparatus.";

RL J. Bacteriol. 180:3062-3069(1998).

CC -I- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA. PROBABLY ACTS AS A PHYSICAL BARRIER TO PARASITES AND LYtic ENZYMES.

CC -I- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-LAYER WITH HEXAGONAL SYMMETRY. SECRETED BY AN ABC TRANSPORTER (TYPE I) SECRETION APPARATUS.

CC -I- MISCELLANEOUS: THE CALCIUM BINDING OF THIS PROTEIN MAY MEDIATE SURFACE ATTACHMENT BY IONIC BRIDGING BETWEEN THE S-LAYER AND A SPECIFIC MEMBRANE-ASSOCIATED OLIGOSACCHARIDE.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF062345; AAC38665.2; --

DR EMBL; AF193063; AAF19365.1; --

DR EMBL; AE005779; AAK22991.1; ALT_INIT.

DR PIR; A48995; A48995.

DR HSSP; P22629; LSWC.

DR TIGR; CC1007; --

DR InterPro; IPR001343; Hemlysn_Ca_bind.

DR Pfam; PF00353; hemolysinCbind; 3.

DR PRINTS; PR00313; CBNDNGRPT.

DR Cell wall; S-layer; Calcium-binding; Complete proteome.

KW INIT MET 0

FT INIT MET 0

SQ SEQUENCE 1025 AA; 98001 MW; AD7A326E1363D8AC CRC64;

Query Match 3.8%; Score 178.5; DB 1; Length 1025;

Best Local Similarity 19.2%; Pred. No. 0.11;

Matches 169; Conservative 116; Mismatches 389; Indels 205; Gaps 36;

Qy 2 SNINVIKSNIOAGLNSTKGLKNLYLAIPKDYDPQKGTINDFIKADELGIARLAEEP- 60

Db 166 ANIDYLTAFVRA--NTPPTAAADIDLAV-----KAALIGTILNATVSGIGYATATA 216

Qy 61 ---NHTETAKKSVDT---VNQFLSLTQTGTGIAISATKLEKFLQKHSNKLAKGLDSVENID 114

Db 217 AMINDLSGDGALSTDNAAGVNLFTAYPSSGVs-----GSTLSLTGTGDTL----- 260

Qy 115 RKLKASNVLSLSSFLGTAL-----AGIELDSLIIKKG--DAAPDALAKASIDLI 162

Db 261 --TGTANNDDTFVAGEVAGAAATLVGDTLGGAGTDVLNWNVQAAAVTALPTGVITSGIETM 318

Qy 163 NEIIG-----NLQSQTQTTEAFESQLAKGLSTISQAKGFSNIGNKLQNLNF-----SKTN 212

Db 319 NVTSGRAITLNTSSGVGTGLTALNTNTSGRAQTVTAG-----QNLTAATAQAANN 370

Qy 213 LGLE---IIFGLLSGTSAGPALADKNASTKKVAAGFELSN-----Q-----Q 251

Db 371 VAVDGGANVTVASTGVTGTTTGVGANSASGTVSVSANSSTTTTGAIAVTGGTAVTVAQ 430

Qy 252 VIGNVTKAISSYVLAQRVAAGLSTTGAVALIITSSIMLAISPLAFWNAADKFNHANALDE 311

Db 431 TAGNAVN--TTLTQADVTVTGNSSITAVTVTQTAATAGATVAGRVNGA-----VTITDS 483

Qy 312 FAKQFRKFGYDGHLLAEYQYRGVGTTEAS--LTTISTALGAVSAGVSAAGVAGTPIA- 369

Db 484 AAASATTAGKIAVTVLGSP--GAATIDSSALTIVNLSGTGTSIGIGR--GALTATPTAN 538

Qy 370 LLVAGVTGL--ISGIL---EASKQAMFESVANRLQKILEWEKQNGQNYFDKQVDSRYAA 425

Db 539 TLTILNVNGLTTTGAITDSEAAADDGFTTI-----NIAGSTASSTIAS 580


```
Qy 926 SV 927
Db 972 TV 973

RESULT 28
RBP2 PLAVB
ID RBP2 PLAVB STANDARD; PRT; 1251 AA.
AC Q00799;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Reticulocyte binding protein 2 (Fragment).
GN RBP2.
OS Plasmodium vivax (strain Belem).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92315338; PubMed=1617731;
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
RL merozoites."
RL Cell 69:1213-1226(1992).
CC -!- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC -!- HUMAN RETICULOCYTE CELLS.
CC -!- SUBCELLULAR LOCATION: Membrane-bound (Probable).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch)
CC or send an email to license@isb-sib.ch.
CC
CC EMBL; M8098; AAA29744.1; -
KW Malaria; Receptor; Membrane.
FT NON TER 1
FT NON TER 1251
SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;

Query Match 3.7%; Score 173; DB 1; Length 1251;
Best Local Similarity 19.3%; Pred. No. 0.28; Indels 294; Gaps 41;
Matches 186; Conservative 142; Mismatches 341;

Qy 28 AIPKDYDPQGGTLDNDFKAADGELIARLAEEP-----NQFLSLTQTGIAISATKLEKFLQ-----KH 98
Db 18 SIEKAYE-KMGNTLKELEKMDDEKNIKEVEEAIQYKRIPIHDHVNLMNDEVEKSKIVM 76
Qy 61 NHTETAKSVDTV-----NQFLSLTQTGIAISATKLEKFLQ-----KH 98
Db 77 EKIELYKKEIDEIKQKNEYKQGDTSNFFYEQYNSATQ-----SKAKIEQFINIATTKK 131
Qy 99 STNKLAKGLDSVENIDRLKGRASNVLSLSSFLGTALAGIELDSLKKKGDAAPALAKAS 158
Db 132 GTSTSDQINLESIKKEVHK--NL-----QLVQKESNMEEMERKQI 171
Qy 159 IDLINEIIGNLQSQTTEAPSSOLAK-LGSTISQAKGF-SNIGNKQLNLFSTNIGLE 216
Db 172 LSMKDLILNSET-----IAKEISNNTQNALGFRENATKLU-----NKTDELQ 216
Qy 217 IITGLLSGISA-----GPAADKNASTGKVAAGFELSNGVGNVTKAISYVLAQRVA 271
Db 217 RVAAMIEEAKAHKNIDIALEDAQIDT---EVSKIEQINREIMNKKDEIKSYLSEIKEYK 273
Qy 272 GLSTGTGAVALITSSIMLAISPLAFMNAADFNHVALDEFAKFRKGYGDHL--LAE 329
Db 274 DKCTTE-----ISNS-----KRGKDKI-----EFLKPKNEESNKNVINE 311
Qy 330 YORGVGTTIASLITITSTALGAVSAGVSAAGVAVGTPIALLVAGTGLISGILEASK-- 387
```

```
Db 312 INENIRNSEQYLKDIEDAEKQAS-----TKVELFHKHET-TISNIFKESEIL 357
Qy 388 --QAMFESVANRLOGKILEWEKONGQNYFDKGYDSRYAAYLANNLKFLS----- 435
Db 358 GVTESQKKINKAEDIMKEIERHNSBIQTQVKGFQ-----ENLKLMEPHINYNAED 409
Qy 436 ELN--KELEAERVIAITQORWNNNIGELAGITKLGRIKSGKAYADAFEDGKKEVAGSNI 493
Db 410 ELNNDKSTNAKVLJETNLESVKHLSLSEITNIKOGGEKI-----YSKAKDIMQIKATSEN 464
Qy 494 TLDKATGIIDISNNGKKTQALHFTSPLLTAGTESRRLTNGKYKSYINKL-KFGRVKNWQ 552
Db 465 TAE-----KTLKVKDQDSNTVYVNLNQITTRNLI 494
Qy 553 VTDGEASSKLD--FSKVIQORVAETEGTDEIGLIVNAK--AGNDDIFVGQCKMNDGDDGH 608
Db 495 VTEKNRLNGIDSTITNIEGALKESKNGYEIGFLEKLEEIGNRKLVKVDITKGSINSTVGN 554
Qy 609 DRVFSYKDGFGNITVDGTSATAGSYTVNRKV-----ARGDIYHEVVKRQETKVGKRT 663
Db 555 ---FSSLPNPF-----DLNQYDFKNINDYENKWEIYNEF---EGLNKISE 596
Qy 664 TIQ-----YRDYELRV--GYGVQSTDLNLSVEEVIGSFQFNDVFKGSKFNDIFHSGEGDD 716
Db 597 NLRNASENTSNDYNSAKTLRLAQKEKVNLLNKEEANKYLRDVKVKVESPRFIENMKESLD 656
Qy 717 LLDGGAGDDRLFCGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIPKGDGNDLYD 776
Db 657 KINEMIKKEQL-----TVNEGHGNVQOLVE----- 681
Qy 777 GTGNDKLAFAADANISDITMIERTKEGIIIVKRNDSHSGSINIPRWYITSNLQYQSNKTDHKI 836
Db 682 --NIKELVDENNLSILKQATCKNEIQTTH-----STLKNKAKTILGH-V 725
Qy 837 EQLIKGDGSVITSD-QIDKILQDKOGTVITSELK-----KLADENKSKQLSASDIA 888
Db 726 DTSKAVYGIKITPELALTLLGDAK---LKTQELKFESKNVNVVLENNMKNTELDVH 782
Qy 889 SSL 891
Db 783 KNI 785

RESULT 29
ZAPA_PROMI
ID ZAPA_PROMI STANDARD; PRT; 491 AA.
AC Q11137;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Extracellular metalloprotease precursor (BC 3.4.24.-).
GN ZAPA.
OS Proteus mirabilis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Proteus.
OX NCBI_TaxID=584;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BB2000;
RX MEDLINE=96011363; PubMed=7592325;
RA Wassif C., Cheek D., Belas R.;
RT "Molecular analysis of a metalloprotease from Proteus mirabilis."
RL J. Bacteriol. 177:5790-5798(1995).
CC -!- FUNCTION: ONE OF THE VIRULENCE FACTORS PRODUCED DURING SWARMER
CC CELL DIFFERENTIATION OF THE BACTERIA, WHICH SEEMS TO BE ASSOCIATED
CC WITH PATHOGENESIS. THE PROTEASE ACTIVITY IS LIMITED TO IG A1,
CC IG A2, AS WELL AS IG G DEGRADATION.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- ENZYME REGULATION: Ca(2+) increases protease activity.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: THE GLY-RICH REPEATS MAY BE IMPORTANT IN THE
CC EXTRACELLULAR SECRETION OF THIS METALLOPROTEASE.
CC -!- SIMILARITY: Belongs to peptidase family M10B.
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL; U25950; AAA86729.1; ALT_INIT.
DR HSP; P23694; 1SAT.
DR MEROPS; M10.057; -.
DR InterPro; IPR001343; Hemlynn Ca bind.
DR InterPro; IPR006026; Nzn Mtpetpase.
DR InterPro; IPR006025; Zn Mtpetpase.
DR Pfam; PF00353; hemolysinCbind; 3.
DR PRINTS; PR00313; CABINDGRPT.
DR SMART; SM00235; ZmC; 1.
DR PROSITE; PS00142; ZINC PROTEASE; 1.
DR PROSITE; PS00330; HEMOLYSIN CALCIUM; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Zymogen; Repeat; Virulence; Calcium.
FT PROPEP 1 16 POTENTIAL.
FT CHAIN 17 491 EXTRACELLULAR METALLOPROTEASE.
FT METAL 186 186 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 187 187 BY SIMILARITY.
FT METAL 190 190 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 196 196 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 226 226 ZINC (CATALYTIC) (BY SIMILARITY).
FT DOMAIN 343 393 4 X REPEATS, GLY-RICH.
FT REPEAT 343 348 1.
FT REPEAT 361 366 2.
FT REPEAT 379 384 3.
FT REPEAT 388 393 4.
SQ SEQUENCE 491 AA; 54000 MW; 1E6DEBE3F6243A97 CRC64;

Query Match 3.7%; Score 170; DB 1; Length 491;

Best Local Similarity 21.0%; Pred. No. 0.12;
Matches 98; Conservative 75; Mismatches 153; Indels 140; Gaps 22;
QY 495 LDKATGIIDISNSGKKTQALHTSPLL-----TAGTE-SRERLT-NGKY----- 537
DB 6 LKKAAGLSNIDLLDKSGIFNYSTKVLPSFDYDTAGKHAREDSWNGKVGIGQPAEVT 65
QY 538 -----SYINKL-KPGRVKWQVTD-----GRASSKLDPSKVIQVARETEGDEGLIVNAKAG 589
DB 66 YSPPKWEKGFNQGNKNPFYENELQKEHARKSLDAWSDIANKIFTEVA--VGNVDGMKAS 123
QY 590 N--DDIFVGQGRWNIDGGDGHDRVF-----YSKD-----GGFNGITVDGTSATSAGSY 635
DB 124 DVKTDITFG-----NIYDNGTFQATATLPNTYAYGKDLGQAWFSDHYHAGNTTPELGNY 179
QY 636 TVNRKVARGDYHEV-----VKRQETKVGKRTETQYRD----- 669
DB 180 -----GRLLTIHEIGTLGLMHPGYNAGQNVPGYKLSDYAEDSRQYTVMSYDWEYETG 233
QY 670 -----YELRKVGIGYQSTDNLKSVVEVIGSQFN---DVFGKSFEND-----I 708
DB 234 AHFOGAYAGAPLLHDISAQVLYGANTTTTRTGDDVYGFNSNTGIDYATATSDNKLIFS 293
QY 709 FHSGBG-----DILLDGGAGD-----DRLFGGKGNDRLSGD 739
DB 284 WDSGGNDTFDSGFQDQLIDLRAQNSFSDVGLQKNVSIQAVNTIENAIAGFGNDIIHGN 353
QY 740 EGDLLDGGSGDDVLLGAGND-VYFRKGDGNDTLYDGTGNDKDLAFADANISIMIMERT 798
DB 354 DADNTLIGEGDDIIYHSGNNTIY---GGRQDTHLGGTSGNTFIYKE--IADSLVTA 408
QY 799 KEGIVKRNDRSGSINIPRWITSNLQNYQSNKTDHKEQLIGKDG 844
DB 409 DKIMDFK-----TGIDKIDLTSLIQDTFSSKILNFVDFNFTGNAG 447

RESULT 30

OMPA_RICRI
ID _OMPA_RICRI STANDARD; PRT; 2249 AA.
AC P15921;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Outer membrane protein A precursor (190 kDa antigen) (Cell surface
DE antigen) (rOmpA) (rOmp A).
GN OMPA.
OS Rickettsia rickettsii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=783;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R;
RX MEDLINE=90354033; PubMed=2117568;
RA Anderson B.E., McDonald G.A., Jones D.C., Regnery R.L.;
RT "A protective protein antigen of Rickettsia rickettsii has tandemly
RT repeated, near-identical sequences";
RL Infect. Immun. 58:2760-2769(1990).
CC -!- FUNCTION: ELICITS PROTECTIVE IMMUNITY.
CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC -!- S-LAYER WITH HEXAGONAL SYMMETRY.
CC -!- PTM: GLYCOSYLATED (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

EMBL; M31227; AAA26380.1; -.
PIR; A41477; A41477.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR005546; Autotransporter.
DR Pfam; PF03797; Autotransporter; 1.
DR TIGRFAMs; TIGR01414; autotrans_bar1; 3.
KW Antigen; Repeat; Signal; Cell wall; S-layer; Glycoprotein.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 2249 OUTER MEMBRANE PROTEIN A.
FT DOMAIN 212 1180 13 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 212 286 A (TYPE I).
FT REPEAT 287 358 B (TYPE II).
FT REPEAT 359 430 C (TYPE II).
FT REPEAT 431 505 D (TYPE II).
FT REPEAT 506 577 E (TYPE II).
FT REPEAT 578 652 F (TYPE II).
FT REPEAT 653 724 G (TYPE II).
FT REPEAT 725 799 H (TYPE II).
FT REPEAT 800 874 I (TYPE II).
FT REPEAT 875 949 J (TYPE II).
FT REPEAT 950 1021 K (TYPE II).
FT REPEAT 1022 1093 L (TYPE II).
FT REPEAT 1094 1165 M (TYPE II).
FT REPEAT 1166 1180 TYPE I (INCOMPLETE).
SQ SEQUENCE 2249 AA; 224333 MW; A9D646C089DF087 CRC64;
Query Match 3.6%; Score 169.5; DB 1; Length 2249;
Best Local Similarity 20.3%; Pred. No. 0.92;
Matches 190; Conservative 117; Mismatches 338; Indels 291; Gaps 46;
QY 1 MSNINVIKS-NIQAGLNSTKSLGNLYLAIPKDYDPQGGTLN-----DFIK----- 46
DB 359 VGNTNALTAVNVGAGLLQVQGV-----VKANTINLTDNASAVTFNPNVVT 405
QY 47 -AADELGIAR---LAEEPNTETA-----KKSVDTVNQFSLTQTGIAL---SATKL---E 92
DB 406 GAIDNTGNANNGIIVTFTGNSTVTGIDTGNLTALATVNVGAGTATLGGAVIKATTTKLTA 465

```
QY 93 KFLQKSTNKLAKG-LDS-----VENIDRKLKGA-----SNVLSLSSFLGTALAG 137
Db 466 SVLTLTANAVLTGAIINTTGGDNVGVNLNGALSQVTHGNTNSLSTISVAGGTATLG 525
QY 138 --IELDSLKKGDAAPDALAKASIDLINELI-----GNLSOSTOTIEAFSSQLAKLG 187
Db 526 GAVIKATTTKLTDA-----SAVKFTNPVVVTGAIDNTGNANNGIVTFTGNSVTVDIG 579
QY 188 -----STISQAKGSNIGKLNLFNFKTNLGLBIIITGLSGISAGFALADKNASTKKV 242
Db 580 NTSNLSTISVAGGTATLGGAVIKATTTKLTNAASVLT--LTNANAVLTGAIDNTTGGDNV 637
QY 243 AAGFELS---NOVIGNTVKAISSVLAORVAAGLSTT-GAVAALITSSIMLAISPLAFMN 298
Db 638 GV-LNLNGALSQVTDIGNTNSLSTIS--VGAGTATLGGAVIKATTTKLTNAASVAFKN 694
QY 299 -----AADKPNHANALDEPAKQPRKFGYDGHLLAEYQVGVGTIEASLTTI-----STA 347
Db 695 PVVVTGAIDSTGNAN-----NGIVTFTGNSVTVDIGNTNA 730
QY 348 LGAVSAGVSAANVSAVCTPIALLVAGVTGLISGILEASQAMPESVANRLOGKILEWEK 407
Db 731 LATNVAGTATLGGAV--IKATTTKLTNAASVLTLTNANAV-----LTGAI-----DN 777
QY 408 QNGGONYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQORWNNIGELAGITKL 467
Db 778 TTGGDN-----VGVLNNGALSQVTDI-----GNTNSLSTIS-- 810
QY 468 GERIKSGKAVADAFEDGKKVEAGSNITLDAKTGIIDISNSGKKTOALHFTSPLLTAGTE 527
Db 811 ---VGAGTATL-----GGAVIKATTTKLTNAASVLTLTNANAVLTGAVDNT---TGSDN 858
QY 528 SRELTNGKYSYI-----NKLKPRVKNQVTDGGEASSKLDPSKVIQVVAETEGTDEI 580
Db 859 VGVLNNGALSQVTDIGNTNSLSTISVAGGTATLGGAVIKATTTKLTNAASVLTLTNAN 918
QY 581 GLIUNA---KAGNDIFVQGGKMDIGDGDHDFVYSKGGFGNITVDGTSATAGSVTV 637
Db 919 AVLTGAIINTTGGDNV-----GVNLNG-----ALSQVTDIGNTNSLSTISVAGGTATL 968
QY 638 NRKVARGDIYHEVVKROETKVKRRTETIQYRDYELRKVGYGYQSTDLNLSVEEVIGSQFN 697
Db 969 G-----GAVIKATTTKL-----TDAASAV----- 987
QY 698 DVPKGSKF-NDIFHSIGEDDLDCAGGDDRLFGCKGNDRLSGDEGD----- 742
Db 988 -----KFTNPVVVTGAIDNT---GNANNGIVTFTGNSVTVDIGNTNALATNVNAGLL 1038
QY 743 -----DLDDGSGDDVLN-----GGAGNDVYIFRKGNDNTLYDGTG 779
Db 1039 QVQGGVVKANTINLTNANASAVTFTNPVVVTGAIDNTGNANNGIVTF---TGNSTVTGNVG 1095
QY 780 NDLKAFADANISDIMERTKEGIIVKRNDSHSGSINI 815
Db 1096 NTN-ALATNVNAGLLQ--VQGGVVKAN-----TINL 1124

RESULT 31
PMPD CHLMU STANDARD; PRT; 1520 AA.
ID -CHLMU
AC O9PLB0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmpD precursor (Polymorphic membrane
DE protein D).
GN PMPD OR TC0197.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Nigg;
```

```

RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Winn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -1- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
CC ENBL; AE002286; AAP39070.1; -.
DR PIR; A81731; A81731.
DR TIGR; TC0197; -.
DR InterPro; IPR005546; Autotransporter.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; DUF145; 1.
DR TIGRFAMs; TIGR01376; POMP repeat; 15.
DR Outer membrane; Signal; Multigene family; Complete proteome.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 1520 PROBABLE OUTER MEMBRANE PROTEIN PMPD.
SQ SEQUENCE 1520 AA; 162102 MW; 68814405AC79585F CRC64;

Query Match 3.6%; Score 169; DB 1; Length 1520;
Best Local Similarity 19.1%; Pred. No. 0.58;
Matches 222; Conservative 143; Mismatches 386; Indels 412; Gaps 58;

QY 10 NIOAGLNTSKYGLKN---LYLAIPKDYD-----PQKGTINDFIKADELGIARLAEEP 60
Db 17 SVVAALASMSGLSNCSDLY-AVGSADHPAYLIPO-AGLLDHIK---DIFIG----- 65
QY 61 NHTETAKSVDTNQVFLSLTQTGTIAISATKLEFLOK--HS-----TNKLAKGLDSVEN 112
Db 66 -----PKSQDKQYKLIIGEAG-SFQDSNAETLPQKVEHSTLFSVTPPIIVGIDQDQ 119
QY 113 IDRK-----LGKASNVLSLTLSSFLGTALAGIELDSLIIKGDAAPD-----ALAKASI 159
Db 120 VSSQGLVCNPSGDHSEIEFERESFLGIAFLG-----NGSKDGITLTDIKSSLGAALYSSD 175
QY 160 DLINELIIGNLSOSTOTIEAFSSQLAKLSTISQAKGSNIGKLNLFNFKTNLGLLEII 218
Db 176 DLIFERIKGDIELSS-----CSSLPRGGACSAQ-----SILIHDCOGLTVKHCAGVNV- 224
QY 219 TGLLSGISA-----GFLADKNASTGKK---VAAGFELSNOVIGNVTKAISVYLAQR 268
Db 225 -----EGVSADHLFGCGGAFSTSSLSGKSLTWPAG-----DIVVATCDGVCFCFG 272
QY 269 VAAGLSTTTGAVAA-----LITSIMLAISPLAFMNAAD---KFNHA 306
Db 273 NSAQLANGAIAASGKVLVFNVEKKISFTDNOALSGGAISASSISIFQNCALVFKSNLA 332
QY 307 NALDEFAKQPRKFGYD-----DHLAEYQVGV-----GTIEASLTTISALG 349
Db 333 KGVKD-----KCSLGGGALASLESVVLKDNLTGITYEKNQSYSEGGAIFGKDCIEFENRG 386
QY 350 AVSAGVSAANVSAVGTPIALLVAGVTGLISGI--LEASKAMPESVANRLOGKILEWEK 408
Db 387 PVVFRONTAALGG--GAILAQQTVAICGNKSGISFEGSKSP----- 426
QY 409 NGGONYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQORWNNIGELAGITKL 468
Db 427 -GG-----AIACGN--FSSENNSS-----ALGSDISNNLGDISFLRTLCL 463
```


[illegible]

RESULT 33

```

PRTC_ERWCH STANDARD; PRT; 479 AA.
ID PRTC_ERWCH AC P16317;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Secreted protease C precursor (EC 3.4.24.-) (Proc).
GN PRTC.
OS Erwinia chrysanthemi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OC NCBI_TaxID=556;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B374;
RX MEDLINE=91009140; Pubmed=2211614;
RA Deleplaire P., Wandersman C.;
RT "Protein secretion in Gram-negative bacteria. The extracellular
RT metalloprotease B from Erwinia chrysanthemi contains a C-terminal
RT secretion signal analogous to that of Escherichia coli alpha-
RT hemolysin."
RL J. Biol. Chem. 265:17118-17125(1990).
RN [2]
RP SEQUENCE OF 1-59 FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=B374;
RX MEDLINE=89255387; Pubmed=2722818;
RA Deleplaire P., Wandersman C.;
RT "Protease secretion by Erwinia chrysanthemi. Proteases B and C are
RT synthesized and secreted as zymogens without a signal peptide.";
RL J. Biol. Chem. 264:9083-9089(1989).
RC -1- COFACTOR: Binds 1 zinc ion per subunit. Magnesium and calcium
CC increase protease activity significantly above the control value.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DOMAIN: THE GLY-RICH REPEATS MAY BE IMPORTANT IN THE EXTRACELLULAR
CC SECRETION OF THIS METALLOPROTEASE.
CC -1- SIMILARITY: Belongs to peptidase family M10B.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J59229; AAA24860.1; -.
CC EMBL; J04736; AAA24862.1; -.
CC
CC DR
CC

```


1344 TVKAADGKTVYVIGNDGVAVD--PSV--VKGQTFKDGAGLRFYFNKLGQLVTSVSGSWYTAN 1399
709 -----FHSGE---GDDLLDGAGDRLFGGKGNDRLSGD--EGDD----LIDGSGDDVL 754
1400 HDWYIQSGKALGTEQYING---QHLYFKDGQHVKGQLVTSVTDGKRVRYDANSQDOAF 1455
755 NGG--AGNDVYIFRKGDNFTLYDGTGNDK-LAPADANISDIMIERTKEGIIVKRNDHSG 811
1456 NKSVTNKGKTYFF-----GNDGTAQTAGNPKGQTFKDG--SDIRF-YSMEGQLV-----TG 1503
QY 812 SINIPRWYITSNLQN---YQSNKTDHKIEQLIGKSGSYITSQDIDKILQDKDGTVLTSQ 868
Db 1504 S-----GWY--SNAQSQWLKYKNGKVLVTLQTVGSGRQYFDENG1-----QAKGKAVRTSD 1552
QY 869 ELKKLADENKSKLSASDIASSLNKLGVSNALFG 902
Db 1553 GKIRYFDEN-----SGSMITNQWKEVNGRYYYFG 1581

RESULT 36
YMYJB CAEEL STANDARD; PRT; 918 AA.
AC P34487;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein F59B2.12 in chromosome III.
GN F59B2.12.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Connell M., Copey T., Cooper J., Coulson A.,
RA Bonfield J., Burton J., Burton R., Favello A., Fraser A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laissner N.,
RA Lacleille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Suleston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstein L., Wilkinson-Sproat J.,
RA Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; Z11505; CAA77581.1; -
DR PIR; G88545; G88545.
DR WormPep; F59B2.12; CE01024.
KW Hypothetical protein.
SQ SEQUENCE 918 AA; 96560 MW; E464FD86B14945DE CRC64;
Query Match 3.5%; Score 163; DB 1; Length 918;
Best Local Similarity 18.5%; Pred. No. 0.59;
Matches 171; Conservative. 133; Mismatches 358; Indels 262; Gaps 38;
QY 83 GIATSATKLEFKLOKSTNKLAKGLDSVENIDRLKGKASNVLTSLSSFLGTALAGIELDS 142
Db 54 GFAMPKPLDASKAAMPVHSSS-----SHKGHHQSGSSSNTHS-----LTVVVGADGKN 99

DR EMBL; L34155; AAA59483.1; -;
DR EMBL; X85107; CAA59428.1; -;
DR EMBL; X85108; CAA59429.1; -;
DR PIR; A55347; A55347.
DR HSSP; P02468; 1TLE.
DR Genew; HGNC:6483; LAMA3.
DR MIM; 600805; -;
DR MIM; 226700; -;
DR GO; GO:0005604; C:basement membrane; TAS.
DR GO; GO:0008544; P:epidermal differentiation; TAS.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR002049; Laminin G.
DR Pfam; PF00053; laminin_EGF_2.
DR Pfam; PF00054; laminin_G_2.
DR SMART; SM00180; EGF_Lam; 2.
DR SMART; SM00282; LamG; 5.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS01186; EGF 2; 1.
DR PROSITE; PS01248; LAMININ TYPE EGF; 2.
DR PROSITE; PS50025; LAM G DOMAIN; 5.
DR Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal;
KW Alternative splicing; Epidermolysis bullosa.
FT SIGNAL 1 20
FT CHAIN 21 1713
FT DOMAIN 46 201
FT DOMAIN 67 113
FT DOMAIN 114 166
FT DOMAIN 167 185
FT DOMAIN 186 769
FT DOMAIN 770 971
FT DOMAIN 978 1140
FT DOMAIN 1147 1307
FT DOMAIN 1366 1530
FT DOMAIN 1537 1710
FT DOMAIN 231 327
FT DOMAIN 396 548
FT DOMAIN 594 621
FT DOMAIN 702 765
FT DOMAIN 1686 1713
FT DISULFID 67 76
FT DISULFID 69 83
FT DISULFID 86 95
FT DISULFID 98 111
FT DISULFID 114 126
FT DISULFID 116 135
FT DISULFID 137 146
FT DISULFID 149 164
FT DISULFID 202 202
FT DISULFID 205 205
FT SITE 658 660
FT CARBOHYD 542 542
FT CARBOHYD 645 645
FT CARBOHYD 745 745
FT CARBOHYD 882 882
FT CARBOHYD 964 964
FT CARBOHYD 1108 1108
FT CARBOHYD 1131 1131
FT CARBOHYD 1325 1325
FT CARBOHYD 1477 1477
FT CARBOHYD 1667 1667
FT VARSPLIC 1 45
FRPS -> KVSYYGILTYQAKSPALPDGMVLLKPKDVLQ
TGRMSIYBETNTPRDLRHGRVHVVEGNFRHSSRAPV
SREBLMTSLRLADRIQGLYFTETQRLTSEVGLSEASDT
GSGRIALAVETACCPAYAGDSC (in isoform B).
W -> R (IN REF. 2).
ATG -> GMC (IN REF. 2).
M -> K (IN REF. 2).
R -> L (IN REF. 2).
E -> Q (IN REF. 2).

FT CONFLICT 1052 1052 D -> A (IN REF. 2).
FT CONFLICT 1184 1184 G -> A (IN REF. 2).
SQ SEQUENCE 1713 AA; 189304 MW; 45EA9BE1017B60D3 CRC64;
Query Match 3.5%; Score 163; DB 1; Length 1713;
Best Local Similarity 20.0%; Pred. No. 1.4;
Matches 205; Conservative 128; Mismatches 333; Indels 360; Gaps 53;
QY 30 PKDYDPQKG-----GTLDNFIKADELGI-----ARLAEPNHTETAKKS 69
DB 190 PKDSPAECDCDCSCVMTLLNDLATWGEQLRLVKSQLQGLSASAGLLEQRMHMETAK- 248
QY 70 VDTVNOFLSLTGTGIAISATKLE-----KFLQKHS-----TNKLAGK 106
DB 249 -DLRNQLLNY-RSAISNHGSKIIEGLERELTDLNEPEFTELQEKAVNSRKAQTLLNNVRA 306
QY 107 LDSVENIDRKLGA-SNVLSLTSLFGLTALAGIELDSLIIKKGAAPD----- 152
DB 307 TQSAKELDVKIKNVIRNVHILLKQISCTDCEGNNVPS---GDFSREWAQAQRMRELNR 362
QY 153 -----ALAKASID-----LINEI-----IGNLSQSTQTIEAFSSQL-AKL 186
DB 363 RNFGKHLREAEADKRESQLLLNRIRTWKTHOGENNGLANSIRDSLNEYEAKLSDELRL 422
QY 187 GSTISOAKGFSNIG-----NKLQNLNFSK-----TNLGLIEI 218
DB 423 QEAAQAQKQANGLQENGERALGAIQVKEINLSQS-DFTKYLTATDSSLLQTNIALQ-- 479
QY 219 TGLLSGISAGFALADKNASTGKVAAGFELSNOVIGNVTTKAISSYVLAQRVAAGLSTGA 278
DB 480 -----LWEKSQKEYEKLAASLEAREQELSDKVRLELSR-----SAGKT--- 516
QY 279 VAALITSSIMLAISPLAFMNAADKFNHNAIDLDFAKQFR--KFGYDGDHLLAEYQVGVGT 336
DB 517 -----SLVEEAKE--HARSLOELAKOLEIKRNASGDEL----- 548
QY 337 IEASLTITSTALGAVSAGVSAAGVAVGTPFIALLVAGVTGLISGLEASQAMFESVAN 396
DB 549 -----VRCVDAATA-----YENILNAIKAA--EDAAAN 574
QY 397 RLOGKILEWEKQNGQNYFDGYDSRYAAYLANNLKPLSE-----LNKEL 441
DB 575 RAAS-----ASESALQTVIKEDLPKRAKTLSSNSDKLLNEAKMTQKKLKQEVSPALNNLQ 629
QY 442 EAERVIAITQORWDNIGELA---GITK--LGERIKSGKAYADAPEDGKKVAGSNIT- 494
DB 630 QTLNIVTVQKEVIDTNLTTLRLDGLHGIQRGDIDAMISSAKS-----MVRKANDITD 680
QY 495 --LDA---KTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKL----- 543
DB 681 EVLDGLNPQTQDVERIKDTYG-RTQNEDF-KKALTADNSVKNLT-----NKLPLDLWR 731
QY 544 KFORVKNQWTDGEASSKLD-FSKVIQVVAETGTDI EGLIVNAKAG-----ND----- 591
DB 732 KIESINQQLPLGNISDNMDRIELIQOARDAASKVAVPMRFNGKSGVEVRLPNDLEDLK 791
QY 592 -----DIFVGGQKWNIDGGDGHDRVPY--SKDGG---FGNITVDG--TSATEAGSYTVNR 639
DB 792 GYTSLSLFLQRPNSRENGGNTENFMVYLGKNDASRDYIGNAVVDGQLTCVYNLGDREBEL 851
QY 640 KVARGDIYHVVKQETKVKRTETIQRDYRLKVGYGYSQSDN---LKSVEEIVGSQF 696
DB 852 QV-----DQILTKSETKEAVMDRVKQRIYQFARLNYTKGATSSKPKETPGVDMOGRNS 905
QY 697 NDVFKGSKFNIDI FHSGEGLDLDGGAGDRLFGCGKGNDRLSGD-----EGDDLQGGSG 750
DB 906 NTLNLDPENVPVYVG-----GYPPDFKL-----PSRLSPFPYKGCIELDDL-----N 948
QY 751 DDVLNGAGNDVIFP-----RKGDGNDTLYDGTGNDKL---AFADANIS 791
DB 949 ENVLS-----LYNFKKTFNLNTTEVEPCRRRKEESDKNYFEGTGYARVPTQPHAPIPTF 1002
QY 792 DIMIERTKE-GII--VKRNDHSGSINIP-----RWYITSNLQNYQ-----SNKTDHKI 836

```
Db 1003 GQTQTVDRLGLFAENCDRPFISINIEDGKLMVRYKLNSLPKRGVGDAINNRDHSI 1062
Qy 837 BQLIGK 842
Db 1063 QIKIGK 1068

RESULT 38
ID YDBA_ECOLI STANDARD; PRT; 2003 AA.
AC P33666; P76087; P76088; P76856; P76857; P76859;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Hypothetical protein ydbA.
OS YDBA OR B1401/B1405.
GN Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
EX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
EX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kaehimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sanpei G., Seki Y., Sivasubramam S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [3]
RP SEQUENCE OF 464-2003 FROM N.A.
RC STRAIN=K12;
EX MEDLINE=92190338; PubMed=1665988;
RA Mozer I., Glaser P., Danchin A.;
RT "Multiple IS insertion sequences near the replication terminus in
Escherichia coli K-12.";
RL Biochimie 73:1361-1374(1991).
CC -!- SIMILARITY: TO S.TYPHIMURIUM ORF NEAR CVSG (AC P25928).
CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; THE GENE CODING FOR
THIS PROTEIN IS INTERRUPTED BY A HYBRID IS2D/IS30C ELEMENT
BETWEEN AMINO ACIDS 839 AND 840.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000237; AAC74483.1; ALT SEQ.
DR EMBL; AE000237; AAC74487.1; ALT-SEQ.
DR EMBL; D90778; BAAL5009.1; ALT-SEQ.
DR EMBL; D90778; BAAL5880.1; ALT-SEQ.
DR EMBL; D90779; BAAL8881.1; ALT-SEQ.
DR EMBL; X62680; -; NOT ANNOTATED_CDS.
DR EcoGene; EG11307; ydbA.
KW Hypothetical protein; Complete proteome.
```

```
FT CONFLICT 489 489 I -> V (IN REF. 2).
FT CONFLICT 495 495 I -> V (IN REF. 2).
SQ SEQUENCE 2003 AA; 205949 MW; B83A12C8B53220EE CRC64;

Query Match 3.5%; Score 162.5; DB 1; Length 2003;
Best Local Similarity 20.4%; Pred. No. 1.8;
Matches 124; Conservative 75; Mismatches 197; Indels 213; Gaps 28;

Qy 430 NLKFLSELNKELEAEERVIATQORWNNIGELAGITKLGRIKSGRAYADAFEDGKKEA 489
Db 185 NLWQIDEANNTVALEGSADGATKQYNHN-----GELVITG-----DNATVNN 228
Qy 490 GSNITLDAK--TGIIDISNSNGKKTQ-----ALFTSPLTATGATESRELTF-- 533
Db 229 NGKTVYDGDSTG--TEINGNNGKVIQDGLDVSGGHGIDITGDSATVDNKGTTMTVDPE 287
Qy 534 -----NGKSYINKLKFRVKN-----WQVTDGEASKLDFSKVIQVATEGTEDEGLI 583
Db 288 SMGIQIDGDKAIVNNEGESTITNGGTGTQINGDDATANNNGKTTV-----DGKDSGTGE 341
Qy 584 VNAKAGNDIDFVGQGNVNDGG-----DGHDRVF 612
Db 342 IN---GNNGKVIQDGLDVSGGHGIDITGDSATVDNKGTTMTVDPEISIGIQVDGQAVV 398
Qy 613 YSK-----DGGFG-NITVDGTSATEAGSYTVNRKVARGDYIYEVVKRQETKV-GKRTET 664
Db 399 NNEGESAITNGGTGTQINGDDATANNNGKTTVDGKSTG-----TEIAGNNGKV 447
Qy 665 IQYRDYELRKVGYGYQSTDNLKSVE-----EVIGSQFNDVFKGSKFNDIFHSGEG 714
Db 448 IQDGLDVSGGHGIDITGDSATVDNKGTTMTVDPEISIGIQI-----DG 491
Qy 715 DDLLDGAGDDRLEFG-----KGNRLSGDEGDLDDG--GSCDDVLNGAGNDVIFR 766
Db 492 DQAIVNNEGESTITNGGTGTQINGDDATANNNGKTTVDGKSTGKI-----AGNIGIV-- 545
Qy 767 KGGNDTLYDGTGNDKLAFADANISDLMIERTEGIIVKENDHSGSINIPRWIITSNLQN 826
Db 546 NLGSLTIVTGG-----AHGVENIGDGTNNKGDIVV---SDTGSIGV----- 585
Qy 827 YQSNKTDHKLIEQLIGKDGYSITSDQIDKILQDKDGTVTITSQELK----- 871
Db 586 -----LINGEGATV-SNTGDVNVSNATGFSITTSNGSKVSLAGSMQVDFSTG 632
Qy 872 -KLADENKSQKLSASDIASSINKLVGSMW---LFGTANSV-----SSNALQ 913
Db 633 VDLNGNNNSVTLAARDL-----KVVGOKATGINVSGDANTVNTGNVLVDKDTADNAE 687
Qy 914 PITQPTQGI 922
Db 688 YFPDPSVGI 696

RESULT 39
ID TOXB_CLODI STANDARD; PRT; 2366 AA.
AC P18177;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DB Toxin B.
GN TOXB OR TCDB
OS Clostridium difficile.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1496;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI 10463;
RX MEDLINE=90326540; PubMed=2374729;
RA Barroso L.A., Wang S.Z., Phelps C.J., Johnson J.L., Wilkins T.D.;
RT "Nucleotide sequence of Clostridium difficile toxin B gene.";
RL Nucleic Acids Res. 18:4004-4004(1990).
```


RN RP SEQUENCE FROM N.A.
 RC STRAIN=VPI 10463;
 RA von Eichel-Streiber C.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1271-2366 FROM N.A.
 RC STRAIN=VPI 10463; PubMed=1603068;
 RX MEDLINE=92293124; PubMed=1603068;
 RA Eichel-Streiber C., Laufenberg-Feldmann R., Sartingen S., Schulze J.,
 RA Sauerborn M.;
 RT "Comparative sequence analysis of the Clostridium difficile toxins A
 and B";
 RL Mol. Gen. Genet. 233:260-268(1992).
 CC -1- DISEASE: CLOSTRIDIUM DIFFICILE PRODUCES TWO TOXINS, AN ENTEROTOXIN
 CC CALLED A AND CYTOTOXIN B.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC
 CC EMBL; X53138; CAA37298.1; -;
 DR EMBL; X92982; CAA63562.1; -;
 DR EMBL; X60984; CAA43299.1; -;
 DR PIR; A27636; A27636.
 DR PIR; S10317; S10317.
 DR InterPro: IPR002479; CW binding.
 DR Pfam; PF01473; CW binding_1; 17.
 DR Pfam; PF04488; Gly_transf_sug; 1.
 DR Toxin.
 KW TOXIN.
 SQ SEQUENCE 2366 AA; 269709 MW; E1024BD8B8A56ADF CRC64;

 Query Match 3.5%; Score 161; DB 1; Length 2366;
 Best Local Similarity 18.4%; Pred No. 2.6;
 Matches 223; Conservative 164; Mismatches 375; Indels 448; Gaps 55;

 QY 3 NINVIKSIQAGLSTKSLKRLYLAI PKDYDPQKGGTLNDFIKAADELGIARLAEPNH 62
 DB DVDSLSTEIAEAILAKEDIS-----PKSIE-----INLGCNMFSSYSINV 707

 QY 63 TET-----AKSVDTNQL-SLTGTGIAISATKLE-----KFLQKHS--TNK--- 102
 DB EETYPGKLLLVKOKISLMPISQDSIIIVSANQYEVRAINSEGRRELLDHSGEWINKES 767

 QY 103 LAKGLDSVENI-----DRKLGKASNVLTSLSSFL-----GTALAGIELDLSLKKGDAPD 152
 DB IKDISKEYISFNPKENKITVSKNLPESLTLLQEIIRNNSSDIELEEKVNLTECEIN 827

 QY 153 ALAKASIDLINEITGNLSQSTOTTEAFSSQLAKLGSTTSIQAKGFSNIGKNLQNLNF----- 208
 DB VISNIDTQIVEERIEEAKNLSDSINYIKDFELIESISDALCDLKQOQNELEDSHFISFE 887

 QY 209 ---SKNLGLEI-----ITGLSGISAGFALADKNASTGK 240
 DB DISETDEGFSIRFINKETGESIFVETEKIIFSEYANHITEISKIK-TGIFDTVNGKLVK 946

 QY 241 KV-----AAGF-----ELSNQVIGNVTKAISSVLAQRAAGLST---TGA 278
 DB KNLDTTHEVNTLNAAFTIQLSIEYNSKESLSNLVAMKVQVYLAQLFSTGLNITDAK 1006

 QY 279 VAALITSSIMLAISPLAFMNA-----DKFNHANALDEFKAFKRFYGDGHLH--AE 329
 DB VVELVSTALDETIDLLPTLSEGLPIIATTIDGVSIGAAIKELSE-----TSDPLLREQ 1059

 QY 330 YQGVGTTEASLTITSTALGAVSAGVSAAGVSGVTPIALLVAGVTGLISGIL-----EA 385
 DB IEAKTIGIMAVNLTTATTITSSLAGISG--SILLVPLAGISAGIPSLVNLVLRDA 1117

 QY 366 SKQAMF-----ESVANRLQKILE-----WEKQNGGQ 412

Db 1118 TKVDYFKHVSLEVTETGCVFTLLDDKIMMPDQLVISEIDFNNSNIVLCKEIMWEGGSG 1177
 QY 413 NYFDKGYDSRYAA-----YLANNLKFLS-----ELNKL-----EASRVTAITQORW 454
 Db 1178 HTVTDDIDHFFSAPSYTIREPHLSIYDVLEVKQKELDLSKDLMLVLPNAPNRVFA-----W 1232
 QY 455 DNNIGELAGI-----TKLGERIK--SGKAY-----ADA-----FEDGKKVEA 489
 Db 1233 ET--GWTGPGURSLNDGTLLDRIRNYEGEFWRYPAFIADALITLTKPRYD-----T 1295
 QY 490 GSNITLDAKTGIIDISNSGKTKQALHFTSPLTAGTE-SRERL-----TNGKYSYINK 542
 Db 1286 NIRINLDSNT-----RSFIVPIIT--TEYIREKLSYFSGGGTYA-LSL 1327
 QY 543 LKFRGVKNQWVTGEA-----SKLDFSKVIQVAETEGTDEGLIVNAK 587
 Db 1328 SOYNMGINIELSESDVMIIDVNVVRDVTIESKIKKGLIEGILSTLSIEENKILNSH 1387
 QY 588 AGNDIDIFVGQKWNIDGGDGHDRVFS-----SKPSEFYYS--NNLKDVKVITKONVNILTYYL 1532
 Db 1388 EIN---PSGE-----VNGSNGFVSLTFSILEGINAIEVDLLSKSYKLLISGELKILMNS 1440
 QY 615 -----XDGFGN-ITVDGTSATEAGSYTVNRKVARG 644
 Db 1441 NHIQKIDYIGFNSQLKNIPIYSFVDSSEKENGFIINGSTKEGLFVSELDPVVLISKVYMD 1500
 QY 645 DIYHEVVKROETKVGKRTETIQYRVELRKVGQYQSTDNLKSVEEV-----IGSQF 696
 Db 1501 D-----KENDIFHSGEGLDLDGGAGDRLFGCKGNDRLSGD----- 739
 QY 697 NDVFKGS-----KENDIFHSGEGLDLDGGAGDRLFGCKGNDRLSGD----- 739
 Db 1533 KDKIKISLTLQDEKTIKLSV-HLDE-----SGVAEILKFMNRKGNNTYSDSLMSFLE 1586
 QY 740 -----EGDDLDDG-----SGDDVLNGGAGNDVYI--FRKGDGNDT 773
 Db 1587 SMNIKSIFVNFLOSNIKFIIDANFIISGTTTSIGQFEPICENDNIQYFIKFTLETNYT 1646
 QY 774 LYDGTGNDKLAFADANISD-----IMERTKEGIIIVKRNHSGSINIPR 817
 Db 1647 LYVGNRQNMIVPEYNDLDDSGDISSTVINFSQKLYGIDSCVNVKVNIPNITDEINIP 1706
 QY 818 WYITSNL-----QNYQSNKTDHKIEQLI-----GKGSYITSQIDKILQDK-- 859
 Db 1707 VYETNNTYPEVIVLDVANIYINEKINVININDLSIRVWSNDGNDFILMSTSENVKSVKIR 1766
 QY 860 -----KDGTV 864
 Db 1767 FVNVFKDKTL 1776

 RESULT 40
 PM21 CHLPN STANDARD; PRT; 1609 AA.
 ID PM21 CHLPN
 AC Q926U5; O9RB58;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable outer membrane protein pmp21 precursor (Polymorphic membrane
 DE protein 21).
 GN PMP21 OR CPN0963 OR CP0897.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia. phila.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CWL029.
 RX MEDLINE=92206606; PubMed=10192398;
 RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";
 RL Nat. Genet. 21:385-389(1999).


```
Db 87 AQTSLGTQTAIGLTERGIVLSAQIDKLLQK---TKAGQALGSAESIVQNAKAKTVLS 143
Qy 126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASALINEIIGNISQSTQTIEAFSSQLAK 185
Db 144 GIQSTLGSVLAGMDLDEAL-QNNNSQHALAKAGLELNTSLIENTANSVKTLDFEGEQISQ 202
Qy 186 LGSTISQAKGSGNIGNKQNLN-NEFKTNLGLIEITLGLSGISAGFALADKNASTGKKVAA 244
Db 203 FGSKLQNIKGLGTGDKLKNIGGLDKAGLGDVJSGLSGATAALVLADKNASTAKKVGAA 262
Qy 245 GFELSNQVIGNVTKAISVYLAQRAAGLSTTGAVALITSSIMLAISPLAFMAADKFN 304
Db 263 GFELANQVGNITKAVSSYLAQRAAGLSTGTPVAALIASTVSLAISPLAFAGIADKFN 322
Qy 305 HANALDEFAKFRKPGYGDHLLABYQRGVGTIEASLTITSTALGAVSAGVSAVAAGSVA 364
Db 323 HAKLSYVAERFKLGYDGDNLLABYQRGVGTIDASTVTAINTALAAIAGGVSAAGSVA 382
Qy 365 GTPIALVAVGTGLSIGLEASKQMPFESVANRLOGKILEWEKONGCONYEDKGYDSRYA 424
Db 383 ASPIALVSGITGVISTILQYKQAMFHFVANKHKNKIVWEKONGHNGFYFENGYDARYL 442
Qy 425 AYLANNLKFLSELNKELEAERVIAITQORWNNNIGELAGITKLGRIKSGKAYADAFEDG 484
Db 443 ANLQDNMFLNLNKLKELQAEVIAITQOQWNNNIGDLAGISRLGEKVLGSKAYADAFEEG 502
Qy 485 KKVEAGSNITLDAKTGIIDI NSNGKKTQALHPTSPLLTACTESRERLTNGKYSYINKLK 544
Db 503 KHIKADKLVLQDSANGIIDVNSGKAKTQHILFRTPLLTPOTEHRERVOTCKYEYITKLN 562
Qy 545 FGRVKNQVOTDGEASSKLDFSKVIQV-----AETEGTDEIGLIVNAKAGNDDIFVQ 597
Db 563 INRVDSWKITDGAASSTFDLTNVVQIGIELDNAGNVTKTKETKIIAKLGGEDNVFVGS 622
Qy 598 GKMNIDGGDHRVPSYKGGFGNITVDGTSATAGSYTVNRKVARGDYIHEVVKRQETK 657
Db 623 GTTEIDGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSHTAL 681
Qy 658 VKRTTETIOYRDYELRKVGYSQSTDNLSKVEEVITGSOFNDFVFGSKFNDIFHSGEGDDL 717
Db 682 VGNREEKTEYR-HSNNQHHAGYTYKDTLKAVEEIIIGTSHNDIFKSGKFNDAFNGDGVDT 740
Qy 718 LDGGAGDRLFGKGNDRLSGDEGDDLDGSGDDVLNGAGNDVYIFRKGUGNDTLYDG 777
Db 741 IYNGDGNDRLFGKGDDDLIDGGNGDDFIDGGKGNLHGGKDDIFVHRKGUGNDIITDS 800
Qy 778 TGNDKLAFADANISIMIERKGIIVKRNHSGSINIPRWY----ITSNLQNYQSNKTD 833
Db 801 DGNDKLSFSDNLKDLTFEKVKHNLVI-TNSKKKVTIQNWFRADFAKEVPNYKATK-D 858
Qy 834 HKIEOLIKDGSYITSDQIDKILQDKDGTVITSQELKKLADENKSKQLSASDITASSLNK 893
Db 859 EKIEEIIQNGERITSKVDDLI--AKNGKITQDELKSKVDNVNELLKHS-KNVTNSLDK 915
Qy 894 LVGSNALFGTANSVSSNALQIPTQPTQGI 922
Db 916 LISSVSAFTSSNDRSNVLVAPTSMLDQSL 944

RESULT 2
B33389
toxin II - Actinobacillus pleuropneumoniae
N:Alternate names: cytolysin II; RTX-toxin II (ApxII)
C:Species: Actinobacillus pleuropneumoniae
C:Date: 09-Mar-1990 #sequence revision 01-Nov-1996 #text_change 18-Jun-1999
C:Accession: B33389; S18853; B43599
R:Chang Y.F.; Young, R.; Struck, D.K.
DNA 8, 635-647, 1989
A:Title: Cloning and characterization of a hemolysin gene from Actinobacillus (Haemophilus)
A:Reference number: A33389; UID:90126233; PMID:2693022
A:Accession: B33389
A:Molecule type: DNA
A:Residues: 1-956 <CHA>
```

```
A:Cross-references: GB:M30602; NID:g141823; PIDN:AAA87232.1; PID:g141825
A:Experimental source: serotype 5
R:Smits, M.A.; Briaire, J.; Jansen, R.; Smith, H.E.; Kamp, E.M.; Gielkens, A.L.J.
submitted to the EMBL Data Library, July 1991.
A:Description: Cytolysins of Actinobacillus pleuropneumoniae serotype 9.
A:Reference number: S18852
A:Accession: S18853
A:Molecule type: DNA
A:Residues: 1-956 <SMI>
A:Cross-references: EMBL:XG1111; NID:g38939; PIDN:CAA43423.1; PID:g38941
R:Smits, M.A.; Briaire, J.; Jansen, R.; Smith, H.E.; Kamp, E.M.; Gielkens, A.L.J.
Infect. Immun. 59, 4497-4504, 1991
A:Title: Cytolysins of Actinobacillus pleuropneumoniae serotype 9.
A:Reference number: A43599; UID:92040145; PMID:1937809
A:Accession: B43599
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-27;948-956 <SM2>
A:Cross-references: GB:XG1111; NID:g38939
C:Comment: This organism causes porcine pleuropneumonia.
C:Genetics:
A:Gene: apxIIA; appA; clyIIA
C:Function:
A:Description: attacks blood cell membranes and causes cell lysis
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F:243-787/Domain: hemolysin A homology <HLYA>
F:719-801/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVYF]-X)
F:557/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match 48.1%; Score 2236.5; DB 1; Length 956;
Best Local Similarity 47.7%; Pred. No. 2.8e-107; Indels 53; Gaps 15;
Matches 453; Conservative 188; Mismatches 255;

Qy 1 MSNINV--IKSNIQAQL-----NSTKSGLNKLYLAIPKQYDPQ 36
Db 1 MSKITLSSLSLQOGLKNGKLNQAGTTLKNGLTQTHSLQNGAKKLIYIPQYDSG 60
Qy 37 KGGTLNDFIKADELGLIARLAEPNHTETAKSVDTVNFQFLSLTQTGTIAISATKLEFLQ 96
Db 61 QGNGVQDLVKAANDLGIIEVWREERSNLDIAKTSFDTTQTKILGFTDGRGIVLFAPOLDNLLK 120
Qy 97 KHSNTNKLAKGLDSVENIDRKLGRKASNVLSLSSPLGTALAGIELDSLIIKKGDAAPDALAK 156
Db 121 KNP--KIGNTLGSSASSISQNIKANVTLVGIIQSLGVSGLVNLNELLQNDKPNQLELAK 178
Qy 157 ASIDLINIEIIGNLSQSTQTIEAFSSQLAKLSTISQAKGFSNIGNKQNLN-NFSKTNLGL 215
Db 179 AGLELTWELVGNITASSVQTVDAPEAQISKLGSHLQNVKGLGSLNKLQNLPLDGLKASLGL 238
Qy 216 EITIGLLSGISAGPALADKNASTCKKVAAGFELSNOVIGNVTKAISSVLAQRAAGLST 275
Db 239 DIISGLLSGASAGIILADKEASTEKKAAAGVEFANQIIGNVTKAIVSSYILAQRAVAGLSS 298
Qy 276 TGVAALITSSIMLAISPLAFMAADKFNHANALDEFAKFRKPGYGDHLLABYQRGV 335
Db 299 TGPVAAIIASTVALLAVSPLSFLNVADRFKQADLKISYSERFQKLGCGDRLADFHRETG 358
Qy 336 TIEASLTITSTALGAVSAGVSAVAAGSAGVTPIALLVAGVTGLISGILEASKQAMFESVA 395
Db 359 TIDASVTITNTALAAISGGVGNASAGSLVGPVALLVAGVTGLITLITILEYSKQAMFEHVA 418
Qy 396 NRLOGKILEWEKONGQNYFDKGYDSRYAAYVLANNLKFLSELNKELEAERVIAITQORWD 455
Db 419 NKVHDIRIVEWEKKH-KNKNYFEQGYDSRHLADLQDNMKFLINLNKELOAERVIAITQORWD 477
Qy 456 NNIGELAGITKLGRIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQAL 515
Db 478 NQIGDLAAISRRTDKISGKAYADAFEEGQHQSDSSVQLDNKNGIINISNTN-RKQTSV 536
Qy 516 HFTSPLLTACTESRERLTNGKYSYINKLKFRGVKNQVOTDGEASSKLDFSKVIQV--- 572
Db 537 LFRTPLLTPGEENREROEGKNSYITKHLIQRVDSWTVTDGASSSSVDFTNVVQRIAVKF 596
```

Qy 573 -----ETEGTDRIGLIVNAKAGNDIFVCGKMNIDGGGHDHVRVFSKXGGGFGNITVD 625
Db 597 DDAGNIIESKDKT-----IIANIAGANDNVFVGSSTTVIDGGDGHDRVHSR-GEYGAIVD 652
Qy 626 GTSATEAGSYVNRKVGDIYHVRKQETKVKRTETIYQYRDYELRKVGYGYSQSTDNL 685
Db 653 ATAETKGSYSVKYVGSKALHETIATHQTNVGNREEKIEYR-REDRPHITGYTVTDSL 711
Qy 686 KSVBEVITGSQPNVFKSGKFNDFIHSGEEDDLDDGGAGDRLRFGGKGNDRLSGDEGDDL 745
Db 712 KSVBEIIGSQPNDFIKSGQFDFVPHGGNGVDITDGNDDHDFGAGDDVDIDGNGNFFL 771
Qy 746 DGGSDVDVLANGAGNDVYIFRKGNDTLYDGTGNDKLAFAADANISIMILERTKEGIIVK 805
Db 772 VGGTNDIISGKNDIYVHKTDGNDSTIDSGQDQKLAFLSDVNLKDLTKFKVDSSLEI- 830
Qy 806 RNDHSGSINTPRWY----ITSNLQYQSNKTDHRIEQLIGKGSYIITSQDIDKILQPKD 861
Db 831 INQKGVKRGWFELEDDLASTVANYKAT-NDRKIEIICKGGERITSEQVQDKLI--KEG 887
Qy 862 GTVITSQELKLADEKSKQSASDIASSLNKLVGSMALFGTANSVSN 910
Db 888 NNQISAEALSKVNDYNTSK-DRQNVNSLAKLISSVGSFPTSSSDFRNN 935

RESULT 3
A35254
leukotoxin A - Pasteurella haemolytica (serotype T10)
N;Alternate names: lktA protein
C;Species: Pasteurella haemolytica
C;Date: 10-Aug-1990 #sequence revision 15-Nov-1996 #text_change 18-Jun-1999
C;Accession: S37145; A35254; S34237; S34235
R;Lainson, A.F.; Aitchison, K.; Donachie, W.
submitted to the EMBL Data Library, September 1993
A;Description: DNA sequence of the leukotoxin A gene from P. haemolytica T10 serotype.
A;Reference number: S37145
A;Accession: S37145
A;Molecule type: DNA
A;Residues: 1-955 <LA>
A;Cross-references: EMBL:Z26247; NID:g400424; PIDN:CAAB1206.1; PID:g400425
R;Highlander, S.K.; Engler, M.J.; Weinstein, G.M.
J. Bacteriol. 172, 2343-2350, 1990
A;Title: Secretion and expression of the Pasteurella haemolytica leukotoxin.
A;Reference number: A35254; MUID:90236888; PMID:2185213
A;Accession: A35254
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 950-955 <HIG>
A;Cross-references: GB:M24197; GB:M34943; GB:M34944
R;Lainson, A.F.; Aitchison, K.D.; Donachie, W.
submitted to the EMBL Data Library, June 1993
A;Description: DNA sequence of the carboxy terminal end of leukotoxin A from the T3 sero
A;Reference number: S34235
A;Accession: S34237
A;Molecule type: DNA
A;Residues: 745-955 <LA2>
A;Cross-references: EMBL:Z22884; NID:g311828; PIDN:CAAB0498.1; PID:g311829
A;Experimental source: serotype T3
A;Accession: S34235
A;Molecule type: DNA
A;Residues: 723-955 <LA3>
A;Cross-references: EMBL:Z22887; NID:g311824; PIDN:CAAB0501.1; PID:g311825
A;Experimental source: serotype T10
C;Function:
A;Description: attacks cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;240-786/Domain: hemolysin A homology <HLA>
F;718-809/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVYF]-X)
F;718-726/Region: repeat
F;727-735/Region: repeat
F;736-744/Region: repeat

F;745-753/Region: repeat
F;754-762/Region: repeat
F;763-771/Region: repeat
F;772-780/Region: repeat
F;781-789/Region: repeat
F;792-800/Region: repeat
F;801-809/Region: repeat
F;801-809/Region: repeat
F;556/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match 48.1%; Score 2236; DB 1; Length 955;
Best Local Similarity 48.2%; Pred. No. 3e-107;
Matches 458; Conservative 178; Mismatches 271; Indels 44; Gaps 17;

Qy 5 NVLKSNIQA---GLN-----STKSGLNVLAIAPKD--YDPQKGGTLNDFTKAA 48
Db 12 NGRSTLTATRGLENRAGSLTQAGQTLKNGAKIILYIPKDYKDSGSGNGLQDLVKAA 71
Qy 49 DELGIARLAEPNHTETAKKSVDTVNFQFLSLTGTGIAISATKLEKFLQKHSNTKAKGLD 108
Db 72 EELGIEVQKEEGNDIAKQATSLCTIQNVLGLTERGIVLSAPQLDKLQK---NKVGQALG 128
Qy 109 SVENIDRKLGKASNVLSLSSFLGTALAGIELDSLKKGDAAADALAKASIDLINEIGN 168
Db 129 SSESTAQNFSAQTVLSGVQGSRTVLAGMDLDEAL-QNESDQLTLAKAGLELTNSLIEN 187
Qy 169 LSOSTOTIEAFSSQALAKLGSTISOAKGFSNIGNKLNQNL-NFSKTNLGLLEIITGLLSGISA 227
Db 188 IANSVQTLDAFSEQISQFGSKLQNVKGLNGLDKLNIGGLDKAGLGLDVKSRLLSGATA 247
Qy 228 GFALADKNASTGKVAAGFELSNOVIGNVTKAISSVLAQVAAAGLSTTGGAVALITSSI 287
Db 248 ALVLADKASTAKKVGAGFELANQVVGNTKAVSSVILAQRVAAAGLSSTGPVAALIASTV 307
Qy 288 MLAIPIAPFMAADKFNHANALDEPAKQPKFGYDGDHLLAEYQRGVGTIEASLTISTA 347
Db 308 AVAISPLSFAGIADKDFRAKSENYAERFKLGYEGDSLAEYQHGCTGTTDASVTAINTA 367
Qy 348 LGAVSAGVSAAGVSGAVGTPIALLVAGVTGLIGSILEASKOAMFESVANLOGKILEWEK 407
Db 368 LAAIAGVSAAGVSGVVASPIALLVSGITGVITLIQYKQAMFEHVANKHKNIVEWEK 427
Qy 408 QNGGQNYFDKGYDSRYAAVLAANNLKFELSELNKELEAEVIAITQQRWDNNIGELAGITKL 467
Db 428 NNGKGVFENGVDARYLANLQDNMFLNKLQAEVIAITQQQWDSNIGLAGISRL 487
Qy 468 GERIKSGAYADAFEDGKKVEAGSNITLDAKTGIIISNSNGKKTQALHTFISPLLTAGTE 527
Db 488 GEKVLGKAYVDAFEEGQHLKADKLVQLDSAKGIIDVTNTGEAKTOHILFRTPLLTGTE 547
Qy 528 SRERLTNGKYSYINKLFGVRKQWQVTDGEASSKLPDFSKVIOQVIA-----ETEGTDEI 580
Db 548 KRERVQTKYEYITKLHNKRVDSWQDKGAASSTFDLTNVVQIRIGVELDHAENVIKTKET 607
Qy 581 GLIVNAKAGNDIDFVCGKMNIDGGGHDHVRVFSKXGGGFGNITVDGTSATEAGSYVNRK 640
Db 608 KIVATLGDGDDNVFVSGTTEIDGEGYDRVHSR-GNYGALTIDATKETEQSGSYTVNRF 666
Qy 641 VARGDIYHEVVKQETKVKRTEIYQYRDYELRKVGYSQSTDNLSKSVESVIGSQFNDVF 700
Db 667 VESGKALHEGTSTHTALVGNREEKIEYR-HSNQHAGYTYKDTLKAVEIIGTSHNDIF 725
Qy 701 KGSKFNDIFHSGEGDDLDDGGAGDRLRFGGKGNDRLSGDEGDDLDDGGSGDDVINGAGN 760
Db 726 KGSKFNDIAFNGGGOVDITDGNDRDLFGGKGDDIDGGNGDDFDIDGKGNDDLHGGKGD 785
Qy 761 DVYIFRKGNDTLYDGTGNDKLAFAADANISIMILERTKEGIIVKRNHSGSINIPRWY- 819
Db 786 DIFVHRGCGDNDSTITESGNDKLSFSDSNLKDITFEKVNHLVI-TNTKQEKVTIQWFR 844
Qy 820 ---ITSNLQYQSNKTDHRIEQLIGKGSYIITSQDIDKILQPKDGTGTVITSQELKLADE 876
Db 845 EABFAKTIQNYVATR-DDKIEIIGQNGERITSQVDELIE--KGNGKIAQSELTKVVDN 901

Qy 877 NKSKLASDIASSLNKLVGSMALFGTANSVSSNALQPIPTOQTGILAPSV 927
Db 902 YOLLKYS-RDASNSLDKLSSASFTSND-SRNVL-----ASPT-SMLDPSL 946

RESULT 4
A43834
toxin II - Actinobacillus suis
N:Alternate names: aspa protein; cytolysin II; RTX-toxin II
C:Species: Actinobacillus suis
C>Date: 31-Dec-1993 #sequence_revision 08-Nov-1996 #text_change 05-Dec-1998
C:Accession: A43834
R:Burrows, L.L.; Lo, R.Y.
Infect. Immun. 60, 2166-2173, 1992
A:Title: Molecular characterization of an RTX toxin determinant from Actinobacillus suis
A:Reference number: A43834; MUID:92267623; PMID:1587585
A:Accession: A43834
A:Molecule type: DNA
A:Residues: 1-956 <BUR>
A:Experimental source: isolate 3714
A:Note: sequence extracted from NCBI backbone (NCBIN:104212, NCBI:P:104211)
C:Comment: This organism causes acute fatal septicemia in young pigs.
C:Function:
A:Description: attacks cell membranes and causes cell lysis
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F:243-787/Domain: hemolysin A homology <HLXA>
F:719-801/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVYIF]-X)
F:557/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match 47.5%; Score 2205.5; DB 1; Length 956;
Best Local Similarity 47.5%; Pred. No. 1.1e-105; Indels 51; Gaps 15;
Matches 450; Conservative 187; Mismatches 260;

Qy 1 MSNINV--IKSNIQAGL-----NSTKSLGNLYLAIPKDYDPQ 36
Db 1 MSKITLSSLSKSLQGLKNGKNLQAGTTLKNGLTQTGHSILQNGAKKLLVYIPQYDSG 60

Qy 37 KGGTUNDRIKAEDEIGIARLAEHPNHTETAKSVTVTVNQFLSLTGTGTAISATKLEKFLQ 96
Db 61 QCGVQDLVKAANDLIGIEVWEERSNLDAIKTSFDTQTKILGFTDRGIVLFAPOKLNLLK 120

Qy 97 KHSYTNKLAGLDSVNDIRKLGKASNLVSLSSPLGTALAGIELDLSLKKGDAAPDALAK 156
Db 121 KNP--KIGNTLGSSISQNIQKANTVLGGISLGSVLGNLNLQNDKDPQLEAK 178

Qy 157 ASIDLINEIIGNLSQSTQTIKAFSSQLAKLSTISQAKGFNIGNKLQNL-NFSKTNLGL 215
Db 179 AGLELTNELVGNIASSVQTVDAFAEQISKLSHLQNVKGLGLSNKLQNLPLDGLKASLGL 238

Qy 216 EIITGLLSGISAGFALADKNASTGKKVAGFELSNOVIGNTVKALSSVYLAQORVAAGLST 275
Db 239 DIISGLLSGASAGLTLADKASTEKKAAGVEFANQIIGNTVKAVSSYILAQORVASCILSS 298

Qy 276 TGAAVALLTSSIMLAISLAPFNVAADKFNHANALDEFAKQFKKCYGCDGHLAEYQRCVG 335
Db 299 TGPVAALITASTVAVSLFSLFVNADKFKQDLINKYSERFQKGYDGRDLADFHRETG 358

Qy 336 TIEASLTTITSTALGAVSAAVGSVAGTPIALLVAGVTGLISGILEASKQAMFESVA 395
Db 359 TIDASVTNTALAAISGGVGAASAGSLVGAAPVALLVAGVTGLITLILEYSKQAMFEHVA 418

Qy 396 NPLQSKILEWEKQNGQYFDKGYDSRYAAVYANLKLFLSELNLELEAEVIAITQQRWD 455
Db 419 NKVDRIVVEWEKKH-KNRYFEQGYDSRHLADIQNMKFLINLNKELQAEVVAITQQRWD 477

Qy 456 NNIGLAGITTKLGERIKSGKAYADAFEGKKVEAGSNITLDAKTGIIIDISNNGKKTQAL 515
Db 478 NQIGLALAIKRTDKISSGKAYADAFEGQHQSYDSVOLQNKNGIINISNTN-RKTSV 536

Qy 516 HFTSPLLTAGTSRRLNNGKYSYINKLFGKRGKQKQVTDGEASSKDLFSKVIQKVAETE 575
Db 537 LPRTPLLTGEBNRRRIQEGKNSITTKLHIQRVDSWTVTVGDASSSVDFTNVQR--EAV 594

Qy 576 GTDEIGLIVNAK-----AGNDIDFVGCKMNIIDGGDGHDRVFFYSKDGFGNITVDG 626
Db 595 KFDAGNIIIESKOTKIIIANIGAGNDNVVSSSTTVIDGGDGHDRVHYSR-GEYALVIDA 653

Qy 627 TSATEAGSYTVNRKVARGDIYHEVVKQETKVGKRTETIOYRDYELRKVGYGQSTDNLK 686
Db 654 TAETEKGSYVKRYVGDASKALHETIATHQTVNGNREKIEYR-REDDRFHTGYTVDLSLK 712

Qy 687 SVEVIGSQNDVFKSKFNDI FHSBGDDLLDGGAGDRLFGGKGNDRLSGDEGDDLLD 746
Db 713 SVEEIIQSQNFIDPKGSQFDDVFEHGGNGVDITDNGDDHLFGGAGDDVIDGGNGNFLV 772

Qy 747 GSGDDVLNGAGNDVYIFPKGDNNDLYDGTGNDKLAFADANISDTMIERTKEGIIVKR 806
Db 773 GGTGNDIISGGKNDIYVHKTDGNDISITDSGGQDKLAFSDVNLKDLTFKKVDSLSRI-I 831

Qy 807 NDHSGSINIPRWY----ITSNLQYQSNKTDHKEQLIGKDGSVITSQDQDKILQDKDG 862
Db 832 NQGEKVRIGNWFLEDDLASTVANYKAT-NDRKIEEIIHGGERITSEQVDKLI--KEGN 888

Qy 863 TVITSQELKLADENKSKLSASISLNKLKLVGSMALFGTANSVSSN 910
Db 889 NQISAEALSKVNDYNTSK-DRQNVNSLAKLISVSGFTSSSDFRNN 935

RESULT 5
B49219
toxin III - Actinobacillus pleuropneumoniae (serotype 8)
N:Alternate names: RTX-toxin III (ApXIIIA)
C:Species: Actinobacillus pleuropneumoniae
C>Date: 19-Dec-1993 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C:Accession: B49219; S48043; S29958
F:Jansen, R.; Briaire, J.; Kamp, E.M.; Gielkens, A.L.; Smits, M.A.
Infect. Immun. 61, 947-954, 1993
A:Title: Cloning and characterization of the Actinobacillus pleuropneumoniae RTX-toxin I
A:Reference number: A49219; MUID:93162836; PMID:8432615
A:Accession: B49219
A:Molecule type: DNA
A:Residues: 1-1052 <JANI>
A:Cross-references: EMBL:X68815; NID:G38956; PIDN:CAA48711.1; PID:G38958
A:Experimental source: strain 405, serotype 8
A:Note: sequence extracted from NCBI backbone (NCBIN:125168, NCBI:P:125170)
F:Jansen, R.; Briaire, J.; van Geel, A.B.M.; Kamp, E.M.; Gielkens, A.L.J.; Smits, M.A.
Infect. Immun. 62, 4411-4418, 1994
A:Title: Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (ApX)
A:Reference number: S48042; MUID:95012630; PMID:7927703
A:Accession: S48043
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1052 <JANI>
A:Cross-references: EMBL:X80055; NID:G558150; PIDN:CAA56358.1; PID:G558152
A:Experimental source: strain 405, serotype 8
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Comment: This organism causes porcine pleuropneumonia.
C:Genetics:
A:Gene: apXIIIA
C:Function:
A:Description: lyses lung macrophages
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: calcium binding; cytolysis; exotoxin; lipoprotein; tandem repeat; thiolester
F:254-804/Domain: hemolysin A homology <HLXA>
F:736-862/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVYIF]-X)
F:571/702/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match 42.1%; Score 1955.5; DB 1; Length 1052;
Best Local Similarity 42.6%; Pred. No. 9.1e-93;
Matches 425; Conservative 166; Mismatches 287; Indels 119; Gaps 19;

Qy 8 KSNITQAGLNSTKSLK-----NLYLAIPKDYDPQKGTLDNDFI 45
Db 19 KRQAKGYDVTNKGLQYGVSAQKALQAAGKAVQKGNKLVLPKPEYDGVSGVGNFFDLV 78

C:Function:

A;Description: attacks blood cell membranes and causes cell lysis
 C;Superfamily: hemolysin A; hemolysin A homology
 C;Keywords: calcium binding; cytolysins; exotoxin; hemolysis; lipoprotein; tandem repeat;
 F;246-791/Domains: hemolysin A homology <HLA>
 F;723-891/Region: 9-residue repeats [G-G-X-G-[DN]-D-X-[LVYF]-X)
 F;563,689/Binding site: palmitate (Lys) (covalent) #status experimental

| | | | | | | |
|-----------------------|--------|--------------|----------|------------|--------|----------------------|
| Query Match | 39.7% | Score | 1845; | DB 1: | Length | 1023; |
| Best Local Similarity | 41.8%; | Pred. No. | 4.2e-87; | | | |
| Matches | 421; | Conservative | 156; | Mismatches | 311; | Indels 118; Gaps 23; |

| | | | |
|----|-----|--|-----|
| Qy | 8 | KSNIQAGLNSKSGKLNLYLAIPKDYDPQKGGTGLNDFPKAADELGLIARLAEPNHTETAK | 67 |
| Db | 32 | KDALKKAEEQTRNAGNRLLILIPDYKGO--GSSLNDLVRTADELGIEVOYDEKNGTAITK | 90 |
| Qy | 68 | KSDVTNVQFSLTGTGIAISATKLEKFLQKH--STNKLAAGLDSVENIDRKLKASNVLS | 125 |
| Db | 91 | QVFGTAEKLGILTERGVTFIPAPQLDKLQYKQAGNKLG--GSAENIGDNLGRKAGSVLS | 147 |
| Qy | 126 | TLSSFGLTALAGIELDSLKK-----GDAPPALAKASIDLINEIIGNLSQSTQITIEPSS | 181 |
| Db | 148 | TFQNFLGTALSSMKIDELIKQKGGNVSSELAKASIELINQLV-PTAASLNNVNSFSQ | 206 |
| Qy | 182 | QLAKLGSTISQAKFSNIGNKLQNL-NFSKTNLGLIITGLLSGTSAGFALADKNASTGK | 241 |
| Db | 207 | QLNKLGSVLSNTHLNGVGNKLQNLPLNDNTGAGDVTVSGTLSAISASFLUSNADADTGT | 266 |
| Qy | 241 | KVAAGFELSNQIVGNVTKAISYVYLAQRVAAGLSTTGAAVALITSSIMLAISPLAFMNA | 300 |
| Db | 267 | KAAAGVELTTKVLGVNGKISQYIIAQRAAAGLSTSAAGAGLIASVVTLAISPLSLIA | 326 |
| Qy | 301 | KDFHNALDFAPKQRFKGYDGDHLLAEBYORGVTTEASLTITSTALCAVSAGVSAAV | 360 |
| Db | 327 | KDFRANKIEESYORFKGLYDGDLSLAAFFHETGAIDASITRSTVLVASSGISAAAT | 386 |
| Qy | 361 | GSAVGTPTALLAVAGVTGLISGILEASKQAMPESVANRLOQKILBEWKONGQONTFKGYD | 420 |
| Db | 387 | TSLVGAPVSALVGATGIIISGILEASKQAMPEHVASKWADVIAWEKKH-GKNTFENGYD | 445 |
| Qy | 421 | SRYAAYLANNIKFISELNKELEAEERVIAITQORDDNNNIGELAGITKGLERIKSGKAYADA | 480 |
| Db | 446 | ARHAAFLDENPKILSQYNKEYSVERSVLITQOHWDTLIGELAGVTRNGDKTLSGKSYIDY | 505 |
| Qy | 481 | FEDCKVYEG---SNITLDAKTGLIIDNSNGKKTQALHTSPLLTAGTESRERLNTGK | 536 |
| Db | 506 | YEEGKRLKKEKPEQKQVDFPLKGNIDUSD--KSTLLKKEVTPLLTPGEBIRERROSGK | 563 |
| Qy | 537 | YSYINKLKFRGVQWQVTD-GEASSKLDFSKVIQRVAETEGT-DEIGLIVNAKAGNDIDF | 594 |
| Db | 564 | YEYITELLVKGVDKWTYKGVQDKSVYDYSNLIQHASVGNQYREIRIESHLDGDDKVF | 623 |
| Qy | 595 | VQOGKMNIDGGGHDHVRYPYSK-DGGFGNITVDGTSATGAGSYTVNRKQVARG--IYHEVV | 651 |
| Db | 624 | LSAGSANITYAGKHDVVYDTKDTGY--LTIDGTKATEAGNYTVTR-VLGDGVKVLQEW | 680 |
| Qy | 652 | KROETKVKRTETTOYRDYELARKV-GYCYQSTDNLKSVEEVIGSQFNDVFKGSEFNDIFH | 710 |
| Db | 681 | KEQEVSVGKRTKTYRSYETHINGKNLJETDNLYSVEELIGTTRADKFFGSKFADIFH | 740 |
| Qy | 711 | SEGEDLIDGGAGDDRLF----- | 728 |
| Db | 741 | GADGDDHIEGNDGNDRLYKDGKNDTLSCGNGDDQLYGGDGNDKLIGGAGNYNLNGCGDD | 800 |
| Qy | 729 | -----GGKGNDRLLSGDEGDDLLDGGSGDDVLNGAGNDVYIFRKGDGNDT | 773 |
| Db | 801 | BLQVQNSLAKNVLSSGGKGNDKLYGSEGADLLDGGEGNDLLKGGYGNDIYRSLSGYGH | 860 |
| Qy | 774 | LYDGTG-NDKLAPADANISDIWIERTKEGIIVKRD-----HSGSINI PRWITSNLQ | 825 |
| Db | 861 | IDDDGGKDDKLUSLADIPRDVAFRREGNDLIMYKAEGNVLSIGHKNGITFRKNF----- | 915 |
| Qy | 826 | NYQSNKTDHKIEQLIGKDGSVITSQIDIKLQDKK-----DGTVITSQ----- | 868 |

RESULT 9

A37205
Leukotoxin A - Actinobacillus actinomycetemcomitans
C:Species: Actinobacillus actinomycetemcomitans
C:Date: 16-Sep-1992 #sequence revision 01-Nov-1996 #text_change 31-Mar-2000
C:Accession: A37205; A60768; B34345; A32276; PH0267; PH0266; S17284
R:Kraig, E.; Dailey, T.; Kolodrubetz, D.
Infect. Immun. 58, 920-929, 1990
A:Title: Nucleotide sequence of the leukotoxin gene from Actinobacillus actinomycetemcomitans
A:Reference number: A37205; MUID:90202154; PMID:2318535
A:Accession: A37205
A:Molecule type: DNA
A:Residues: 1-1055 <KRA>
A:Cross-references: GB:X16829; NID:938643; PIDN:CAA34731.1; PID:G38645
A:Note: the authors present evidence that the nucleotide sequence is correct in the vicir
sing nucleotide
R:Kolodrubetz, D.; Dailey, T.; Ebersole, J.; Kraig, E.
Infect. Immun. 57, 1465-1469, 1989
A:Title: Cloning and expression of the leukotoxin gene from Actinobacillus actinomycetemcomitans
A:Reference number: A60768; MUID:89212893; PMID:2707855
A:Accession: A60768
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 297-309, 'Y', 311-364; 434-440, 'KC', 443-474, 'H', 476-489, 'S', 491-493, 'VLK', 497-498
A:Note: this sequence is revised in reference A37205
R:Lally, E.T.; Golub, E.E.; Kieba, I.R.; Taichman, N.S.; Rosenbloom, J.; Rosenbloom, J.C
J. Biol. Chem. 264, 15451-15456, 1989
A:Title: Analysis of the Actinobacillus actinomycetemcomitans leukotoxin gene. Delineation
A:Reference number: A34345; MUID:89359382; PMID:2670940
A:Accession: B34345
A:Molecule type: DNA
A:Residues: 1-239, 'Y', 241-259, 'H', 261-335, 'A', 337-415, 'S', 417-438, 'S', 440-723, 'N', 725-926
PKKRYPLMRCHRPITLTLTQR' <LAL>
A:Cross-references: GB:M27933
R:Lally, E.T.; Kieba, I.R.; Demuth, D.R.; Rosenbloom, J.; Golub, E.E.; Taichman, N.S.; G
Biochem. Biophys. Res. Commun. 159, 256-262, 1989
A:Title: Identification and expression of the Actinobacillus actinomycetemcomitans leukoto
A:Reference number: A32276; MUID:89165963; PMID:2647082
A:Accession: A32276
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 430-438, 'S', 440-476, 'R', 478-506, 'RVRS', 511, 'QSTAINSLNTD', 523-541, 'I', 543, 'PQ
A:Note: this sequence is revised in reference A34345
R:Ohta, H.; Miyagi, A.; Kato, K.; Fukui, K.
submitted to JPID, July 1995
A:Description: Modulation of leukotoxin production by growth rate and bicarbonate in a t
A:Reference number: PH0266
A:Accession: PH0267
A:Molecule type: protein
A:Residues: 17-42 <OHT>
A:Experimental source: strain 301-b
A:Accession: PH0266
A:Molecule type: protein
A:Residues: 2-6, 'L', 8-26 <OHT>
C:Genetics:
A:Gene: ltxA
C:Function:
C:Description: lyses human polymorphonuclear lymphocytes and monocytes
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: cycloolysis; hemolysis; lipoprotein; periplasmic space; tandem repeat; thioles
F:245-790/Domain: hemolysin A homology <HLYA>
F:731-847/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIVF]-X)
F:562,687/Binding site: palmitate (lys) (covalent) #status predicted
Query Match 39.1% Score 1816; DB 1; Length 1055;

Best Local Similarity 40.2%; Pred. No. 1.4e-85;
Matches 413; Conservative 185; Mismatches 299; Indels 130; Gaps 23;

```

QY      3  NINVIKSNIOAGLNSTKSGKLNKLVLAIPKDYDPOKGGTFLNDFIKAADELGAELGARLAESPNNH 62
Db      27  NIDAKEQQLQALDKLGTGKKLYLIPKNY--KKGNLTAIIKAAQKLGIEVTHEGKDG 84

QY      63  TETAKKSVDTVNQFSLSTQTGTGIAISATKLEKFLQ--KHSTNKLAKGLDSDVENIDKRLGKA 120
Db      85  PALTNGLINTGKKLGLTERGLTFAPELDKWIOGNKHLNSVG----STGNLTAKADKV 140

QY      121  SNVLSTSSFLGTALAGIELDSLJK---KGDAAFDA-LAKASIDILINEIGNLSQSQTQI 176
Db      141  QSVLGTTLQAFINTAFSGMDLALIKARQNGKRVNTDVQLAKASLNILINELIGTISITNNV 200

QY      177  EAPSSQIAKLGGSTISOAKGFNSGNKLNQLN----NFSKTNILGLBIIITGLSGISGAFALA 232
Db      201  DTFSGKLNKUGAALQGVKHFGSGDKLNLPKLNGLK---GLGALSGVLSAISAAULLA 257

QY      233  DKNASTGKKVAAQGFELSNQVIGNVTYKTAISSYVLQAQVAAGLSTTGCAAALITSSIMLAIS 292
Db      258  NKDADTAKAAAELTNKVLGNIGKAITQVLIQAQAAAGLSTTGPVAGLIASVLSAIS 317

QY      293  PLAFMNAADKPNHANAIDEPKQPRKGYGDGHHLLAEYQROGVGTIEASLTTISTISTALGAVS 352
Db      318  PLSPLGIKQPDRAARMLEEYSKPEKFGYNGDSLIGQFYKNTGTADAAITTTINTVLSAIA 377

QY      353  AGVSAARAVGSAGVCTPIALLVAGVTGLTSGIILEASKQAMFESVANRLOGLKILEWEKQNGQ 412
Db      378  AGVGAASAGSLVGAPIGLLVSATISLGIILDSAKQAVFEHIANQLADKIKAWENKY-GK 436

QY      413  NYFPKGYSRVAAYLANNLKLPSLNLKELEAEVIAITQORWDNNINIGELAGITKLGBRIK 472
Db      437  NYFENGYDARHSAFLEDSLKLFLNELREKYKTENILSIQOGWDQRIAGELAGITNGDRIQ 496

QY      473  SGKAYADAFEDGKKVEAGSN---ITLDAKTGIIIDISNSNGKTKQALHFTSPLLTAGTES 528
Db      497  SGKAYDVLYKKGEELAKSHDKFTKQILDPIKGNIDLSGIKGSTT--LTFLNPLLTAGKEE 554

QY      529  RERLTNGKYSYINKLKTGRVKNMQVTDGEASSKL-DFSQVLR--VAETEGTDETLGIWNA 586
Db      555  RKTQSGKYEFTLKKVGRTDWKVGVPNSNGVDFSNLIQHAVTRDNKVLLEARLIANL 614

QY      587  KAGNDDIFVGQKQWNI DGGDGHDRVFYSKDGCGFNITVDGTSATEAGSYTVNRKVARGDI 646
Db      615  GAKDDYFVGSSTIVNAGDGYVDVYSK-GRTGALITDGRNATKAGQKVERDLSGTQV 673

QY      647  YHEVVKQETKVKRRTETIQVRDEL--RKVGYGQVSTDNILKSVEEVIGSFQNFDFVFGSK 704
Db      674  LQETVSKQETKRGKVTDLLEVRNYKLDYYNTKGFKAHDELSNVEEIIIGSTLRDKFYGSK 733

QY      705  FNDIFHSGEGDDL-----LDGGAGDRLF----- 728
Db      734  FNDVFHGDGDDLTYGVDGDRLYCDNGNDIEHGGQGNDKLYGGAGNDRLFEGYNNVLD 793

QY      729  -----GGKGNDRLSDEGDDLLDGGSGDDVLNGAGNDVYIFRKGDG 770
Db      794  GGEGDDHLEGGNSDILRGSGSNDKLFQNGQDGLLDGEGDDQLAGGENDIYYRKEYG 853

QY      771  NDTLYDGTGN-DKLAFADANISIMIERTKEGIIIVKRNHDSGSINIPRWYITSN----LQ 825
Db      854  HHTITEHSGDKDKLSLANILNKVDSFERNGNDLLKTNRT-AVTFKGWFSKPNSSAGLD 912

QY      826  NYQ-----SNKTDHKTEQLTGKSGSYITSDQIDKILQDK 859
Db      913  EYQRKLIYAEPKDRARLKRQFELQRGKVDKSLNNKVEEIIIGKQGERITTSQDIDNLPDKS 972

QY      860  KDGTVITSQELKLL-ADENKSQKLSASDIASS-----LNKLVGSMALFCTA-NS 906
Db      973  GNKXTISFQELAGLIKNGKGSLSLMSRSSRSMLTQKSLSNDISRIISATISGFGSSGKA 1032

QY      907  VSSNALQ 913

```

Db 1033 LSASPLQ 1039

RESULT 10

I39643

RTX-toxin I - Actinobacillus pleuropneumoniae

N:Alternate names: Hemolysin Apxi

C:Species: Actinobacillus pleuropneumoniae

C:Date: 19-Jul-1996 #sequence, revision 08-Nov-1996 #text_change 18-Jun-1999

C:Accession: I39643; I39645; S60732; S35781

R:Jansen, R.; Briaire, J.; Kamp, E.M.; Gielkens, A.L.; Smits, M.A.

Infect. Immun. 61, 3688-3695, 1993

A:Title: Structural analysis of the Actinobacillus pleuropneumoniae-RTX-toxin I (Apxi) O

A:Reference number: I39641; MUID:93366425; PMID:8359891

A:Accession: I39643

A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL

A:Molecule type: DNA

A:Residues: 1-1022 <RES>

A:Cross-references: EMBL:X73117; NID:G312897; PIDN:CAA51548.1; PID:G312899

R:Freij, J.; Meier, R.; Gygi, D.; Nicolet, J.

Infect. Immun. 59, 3026-3032, 1991

A:Title: Nucleotide sequence of the hemolysin I gene from Actinobacillus pleuropneumoniae

A:Reference number: S18769; MUID:91348845; PMID:1879928

A:Accession: S18769

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-209, 'AMPYITLA', 218-373, 'R', 375-561, 'Q', 563-686, 'TC', 688-1022 <FRE>

A:Cross-references: EMBL:X52899; NID:G38949; PIDN:CAA37081.1; PID:G38950

R:Freij, J.; Haldemann, A.; Nicolet, J.; Boffini, A.; Prentki, P.

Gene 142, 97-102, 1994

A:Title: Sequence analysis and transcription of the apxi operon (hemolysin I) from Actin

A:Reference number: I39644; MUID:94237497; PMID:8181764

A:Accession: I39645

A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL

A:Molecule type: DNA

A:Residues: 1-209, 'AMPYITLA', 218-373, 'R', 375-561, 'Q', 563-686, 'TC', 688-1022 <RE2>

A:Cross-references: EMBL:X68595; NID:G505568; PIDN:CAA4886.1; PID:G505570

R:Tascon, R.I.; Vazquez-Boland, J.A.; Gutierrez-Martin, C.B.; Rodriguez-Barbosa, I.; Rod

Mol. Microbiol. 14, 207-216, 1994

A:Title: The RTX haemolysins Apxi and ApxII are major virulence factors of the swine pat

A:Reference number: S60731; MUID:95131743; PMID:7830567

A:Accession: S60732

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 604-685 <TAS>

C:Comment: This organism causes porcine pleuropneumonia.

C:Genetics:

A:Gene: apxiA

C:Function:

A:Description: attacks blood cell membranes and causes cell lysis

C:Superfamily: hemolysin A; hemolysin A homology

C:Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat,

F:243-789/Domain: hemolysin A homology <HLXA>

F:721-847/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVYF]-X)

F:560,686/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match 37.9%; Score 1763.5; DB 1; Length 1022;

Best Local Similarity 40.0%; Pred. No. 6.5e-83;

Matches 400; Conservative 181; Mismatches 309; Indels 109; Gaps 26;

Qy 8 KSNITQAGLNSTKSGL-----KNLYLAIPKDYDPKGGTNDPFIKAADELGIARLAEEP 60

Db 21 KSAACGAGALKNGLGQVKGQAGKLIILPKDYQASTGSSLDLVKAAEALGIEVHRSEK 80

Qy 61 NHTETAKSVDTVNOFLSLTQTGTIATSATKLEFKLQKHSNTKLAKGL-DSVENIDRLKLG 119

Db 81 NGTALAKELFGTTEKLGFSGRGIALFAPQDFKLLNKN--OKLUSKLGSGSEALGQRLNK 138

Qy 120 ASNVLTSLSSFLGTALAGIELDSLKK-----GDAAPDALAKASIDLINEIIGLSQSTQT 175

Db 139 TQTALSAQLGFTGTATGMDLSSLRRNRGVEDVSGSELAKAGVDLAQQLVDNIASATGT 198

Qy 176 IEAFSSQLAKLGSITSAQKSGFNIGNKLQNL-NFSKTNLGLGILEITITLLSGISAGFALADK 234

Db 199 VDAFAEQLGKLNALNSTR-LSGLASKLNLPDLISLAGPFGDAVSGILSVSASFILSNK 257
Qy 235 NASTGKKVAAGFELNSQVGNVTKAISVYLAQVAAGLSTTGAAALITSSIMLAISPL 294
Db 258 DADAGTAAAGIETISKILGNIGKVAQYIAQVAAGLSTTAATGGLIGSVVALAISPL 317
Qy 295 AFNNAADKFENANALDEFKQKFGYDGDHLLAEYQYGVGTTEASLTITSTALGAVSAG 354
Db 318 SFLNADKPERAKOLEQYSEKFKFGYEGDSLASFYRETGALEAULTTINSVLASAASG 377
Qy 355 VSAAGVSAVGTPIALVAGVTGLISGILEASKQAMPESVANRLOQKILEWKGQNGQNY 414
Db 378 VGRAATGSLVAPVALVSAITGIISGILDASKQAIFERVATKLANKIDEWKXH-CKNY 436
Qy 415 FDKGYDSRYAAYLANNLKFLSELNKELEABRVIAITTOORWNNIGELAGITKGERIKSG 474
Db 437 FENGYDARHSAFLDTELLSQYNKEYSVRVAITTOORWNNIGELAGITRKGSDTKSG 496
Qy 475 KAYADAFEDGKKVEAG---SNITLDAKTGIIIDISNGKKTQALHFTSPLLTAGTESRE 530
Db 497 KAYVDFEEGKLEKEDPRDKVDFPLEGKIDUSSIN--KTLLKFPVPTVFTAGEIRE 554
Qy 531 RLNTGKYSYINKLFGKRVKQWQVTDGEASSKL-DFSXVIQVAFTEGTDEIGLIVNAKAG 589
Db 555 RKQTKGYEYMTFLVFKCKEKKVVTGVOSHNAIYDTNLIQLAIDKKG-EKQVYTIESHLG 613
Qy 590 --NDDIFVGQKMNIDGGDGHDRVFSK-DGGFNGIIVTDGTSATEAGSYTNRKV-ARGD 645
Db 614 EKNDRIYLSGSSIVYAGNGHDVAYYDKTDGTG--LTFDQSAQAGEYIVTKELKADVK 671
Qy 646 IYHEVVKROETKVKRTETLOYRDELR--KVGYGYOSTNLKSVVEVIGSQFNDVPKGS 703
Db 672 VLKVVKTQDISVSKSEKLEYRDELSPELFGNGIRAKDELHSEVEIISNRKDKFPGS 731
Qy 704 KFNDIFHSGED-----DLLDGGAGDRLRFGKGNDRILSGDEGDDLL 745
Db 732 RPTDIFHGAKGDDEIYNGDGHILYDGDNDVHGGDNDHVLGNGNDRILGKGNFL 791
Qy 746 DGGSGDD-----VLNGAGNDVY-----IFKGDGNDTLVDGTGND----- 781
Db 792 NGDGDDELQVFEGQYNVLGGAGNDILYSGDGTNLPDGGVGNKLYGGLKQIYRYKE 851
Qy 782 -----KLAFADANI-----SPIMERTKEGIIVKXNDHSG-SINIPR 817
Db 852 YGRHIIIEKGGDDDTLLSLDSFKDVGFIKGGDLLVNRKIRGGTLTYHYHEDYNGALTIKD 911
Qy 818 WYITSNLQYOSNKTDKIEOLIGKGSYITSDDIKILQDKDGTVITSQELKKLADEN 877
Db 912 WF--KEGEGQNN---KIEKIVDKGAYVLSQVLTETLAPGRGINFNGLEKLYYGE 965
Qy 878 KSQKLSASDITASSLKLVLGSMALF-GTANSVSSNALQPI 915
Db 966 YN---ALPOLRKDIEQIISSTGAFDGHGKVSVGGGGL 1001

RESULT 11
T00227
hemolysin A toxin protein - Escherichia coli plasmid p0157
C:Species: Escherichia coli
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00227; T42148
R:Makino, K.; Ishii, K.; Yasunaga, T.; Hattori, M.; Yokoyama, K.; Yatsudo, H.C.; Kubota, S.; Shinagawa, H.
DNA Res. 5, 1-9, 1998
A:Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic
A:Reference number: 214127; MUID:98290540; PMID:9628576
A:Accession: T00227
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-998 <MAK>
A:Cross-references: EMBL/AB011549; NID:94589740; PIDN:BAA31774.1; PID:G3337015
A:Experimental source: strain EHEC O157:H7, substrain RIMD 0509952

R:Burland, V.; Shao, Y.; Perna, N.T.; Plunkett, G.; Sofia, H.J.; Blattner, F.R.
Nucleic Acids Res. 26, 4196-4204, 1998
A:Title: The complete DNA sequence and analysis of the large virulence plasmid of Escher
A:Reference number: 220268; MUID:98391744; PMID:9722640
A:Accession: T42148
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-998 <BUR>
A:Cross-references: EMBL/AF074613; PIDN:AAC70116.1
A:Experimental source: strain EDL933; serotype O157:H7
C:Genetics:
A:Gene: hlyA
A:Genome: plasmid p0157
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: cytotoxicity; hemolysis; lipoprotein; toxin
F:233-776/Domain: hemolysin A homology <hlyA>
F:550,675/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match 37.6%; Score 1748; DB 2; Length 998;
Best Local Similarity 40.0%; Pred. No. 3.9e-82;
Matches 391; Conservative 177; Mismatches 308; Indels 102; Gaps 21;
Qy 8 KSNIQAGLSTKSGLNLYLAIPKOYDPQKGGTLNDFPKAABELGIARLAEBPNHTETAK 67
Db 18 KSAFTASSSVRSAGKKILLIPDNYEAQ-GVGINELVKAADBELGIEIHRTERDDTAIAN 76
Qy 68 KSVDTVNFQSLSTQTGIAISATKLEKFLQKHSNTKLAKGI-DSVENIDRLKGLKASNVLS 126
Db 77 QFPGAEEKVVGTLFGRVAFAPQLDKLQKY--QKVGSKTGGTAENVGNLKGAGTVLSA 134
Qy 127 LSSFLGTALAGIELDSLKK---GDAAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQ 182
Db 135 LQNTGTALSGMALDELLRKQREGEDISQNDIAKSSIELINQLVDTVSSINSTVDSFSEQ 194
Qy 183 LAKLGSTISQAKGFSNGKQLNL-NFSKTNLGLIETGLLSGISAGFALADKNASTGKK 241
Db 195 LNQLGSLSSKPLSSVGGKQLNLPGLGDLGVVSGILSVASFSFILGNSDAHTGTK 254
Qy 242 VAGFELSNOVIGNVTKATSSVYLAQVAAGLSTTGAAALITSSIMLAISPLAFNNAAD 301
Db 255 AAAGIELTQVLGNVGKAVSQYILAORMAOGLTAAASAGLITSAVMLAISPLFLAAN 314
Qy 302 KFNHANALDFAKQFRKFGYDGDHLLAEYQYGVGTTEASLTITSTALGAVSAGVSAAVG 361
Db 315 KFERAQLESYSERFKLVNEGALLAAAFHKEGALDAALTTINTVLSVSVAGVSAASA 374
Qy 362 SAVGTPIALVAGVTGLISGILEASKQAMPESVANRLOQKILEWKGQNGQNYFDKYDS 421
Db 375 SLIGAPISMLVSALTGTISGILEASKQAMPESVANRLOQKILEWKGQNGQNYFDKYDS 433
Qy 422 RYAAAYLANNLKFLSELNKELEABRVIAITTOORWNNIGELAGITKGERIKSGKAYADAF 481
Db 434 RHAAFLEDSLSLLADFSRQHAVERAVAITTOOHDEKIGELAGITRNADRSQSGKAYINYL 493
Qy 482 EDGKKEAG---SNITLDAKTGIIIDISNGKKTQALHFTSPLLTAGTSRERLTNGKY 537
Db 494 ENGLLEAQPKFTQVDFPKGTIDL--STGNVSVLTFTPTFTPEEVRERKQSKY 551
Qy 538 SYINKLKFGKRVKQWQVTDGEASSK--LDFSXVIQVAFTEGTDEIGLIVNAKANDDIFV 595
Db 552 EYMTSLTVNGKDTWSV-KGINKHKVYDYSKLQFVEKKNKHQARIISLGDGKDDVYVS 610
Qy 596 GQGWINDGGDGHDRVFSKQDGGFNGITVDTGTSATEAGSYTNRKVARGD--LYHEVVKR 653
Db 611 GAGSSEVFAGEGYDTSYNTK-DVGKLTIDATGASKPGEYIVSKNM-YGVKVLQEVVKE 668
Qy 654 QETKVGKRTETIQRDYELRKVGVYQSTDNLSKVEVIGSQFNDVPKGFNDIFHSG- 712
Db 669 QEVSVGKRTETIQRDYEFETGGIPYDVINDLSHVEELIGKHHDEFKGGKFNDFHGD 728
Qy 713 -----EGDDLDDGGAGDRLRFGKGNDRILSGDEG----- 742
Db 729 GNDYIEGNYGNDRILYDGDGDIYISGGQDDQLFGSGNDKLSGGDGNLTLTGGSGNDELQ 788

Qy 743 -----DLDDGGDDVNLGGAGNDVYIFRKGDNLTLD-GT 778
Db 789 AHGAYNLSGGTGDKLYGGGGIDLLDGGEGNDYLVGGFNDIYVYQNYGHHTIADEGG 848
Qy 779 GNDKLAFADANISDIMITERTKEGIIVKR-----NDHSGSINIPRWYITSNLQNY 827
Db 849 KGDRHLHSDISFDDIAFRVGNLDLIMKAINGVLVSFNESNDVNG-ITFKWFAKD----- 902
Qy 828 QSNKTDHKEIQIKGKGSYITSDQIDKIDKQKGTVTITSQELKKLADENKSQKLSASDI 887
Db 903 ASGADNHLVEITDKGREI---KVDKIPHNNE-----RSGYIKASINASEKNMWNITSV 955
Qy 888 ASSLNKLGVSMALFGTAN 905
Db 956 ANDINKIISVSGFDSGD 973
RESULT 12
I41078
hemolysin - Escherichia coli
C:Species: Escherichia coli
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 21-Jul-2000
C:Accession: I41078
R:Schmidt, H.; Beutin, L.; Karch, H.
Infect. Immun. 63, 1055-1061, 1995
A:Title: Molecular analysis of the plasmid-encoded hemolysin of Escherichia coli O157:H7
A:Reference number: I41077; PMID:7868227
A:Accession: I41078
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-998 <RES>
A:Cross-references: EMBL:X79839; NID:9860924; PIDN:CAA56234.1; PID:94388764
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: lipoprotein
F:233-776/Domain: hemolysin A homology <HLYA>
F:550,675/Binding site: palmitate (Lys) (covalent) #status predicted
Query Match 37.2%; Score 1729; DB 2; Length 998;
Best Local Similarity 39.8%; Pred. No. 3.7e-81;
Matches 389; Conservative 174; Mismatches 313; Indels 102; Gaps 21;
Qy 8 KSNIQAGLNTSGKLNLYLAIPKDYDPQKGTGLNDFPKAADLGIALRLAEENHTEFAK 67
Db 18 KSAFTASSSVRSAGKULLIIPDNYEAQ-GVGINELVKADELGIEIHRTERDDTAIAN 76
Qy 68 KSDVTWVNFSLTQTGIAISATKLEFLQKHSTNKLAKGL-DSVENIDRKLKGSANVLST 126
Db 77 QFFGAERKVVGLTERGVAIFAPQLDKLQKY--QKVGSKIQTAEENVGNLKGAGTVLSA 134
Qy 127 LSSFLGTALAGIELSLIKK-----GDAAPDALAKASIDLINELIIGNLSQSTQTTEAFSSQ 182
Db 135 LQNFITGIALSGMALDELLARKOREGEDIQNDIAKSSIELINQLVDTVSSINSTVDSFSEQ 194
Qy 163 LAKGSTTSQAKGFSGNIGKQLNL-NFSKTNLGIILITGLSGISAGFALAKDAKNASTCKK 241
Db 195 LNLQGSFLSSKPRLSVSGKLNLPDLPLGDLVDVSGILSVASASFILGNSDHTGK 254
Qy 242 VAAGFELSNQVIGNYTKAITSYVLVAQRAAGLSTTGAVAAALITSSIMLAISPLAFMNAAD 301
Db 255 AAAGIELTQVLGNVGKAVSQYILAQRAQGLSTTAASAGLITSAVMLAISPLSFLAAAD 314
Qy 302 KFNHANALDEFAKQFRKFGYDGDHLLAEYQRCVGVGIEASLTIITSLGAVSAGVSAAVG 361
Db 315 KFERAKOLESYSERPKLNYEGDALLAGFHKEGTGAIDAGLTINTVLSSVAGVSAASA 374
Qy 362 SAVGTPIALVAGVTGLISGLEASKQAMFESVANRLOKILEWKGKQNGYFPDKYDS 421
Db 375 SLIGAPIELWSALTGTISGLEASKQAMFESVANRLOKILEWKGKQNGYFPDKYDS 433
Qy 422 RYAAYLANNKPLSELNKELEAEVIAITQQRWDNNIGELAGITKLGRIKSGKAYADAF 481
Db 434 RHAAPLEDSLSLADFQRHQAVERAVAITHQHWDEKIGELAGITRNADRSQSGKPYINYL 493

Qy 482 EDGKKVEAG---SNITLDAKTGIIDISNSGKKTOALHFTSPLLTAGTSRERLRTNGKY 537
Db 494 ENGLLEAEQKPEFTQQVDFQKGTIDL--STGNVSSVLTFITPTFTGEEVERERKQSGKY 551
Qy 538 SYINKLGRVKNQVTVGDEASSK--LDFSKVIQORVAETEGTDBEIGLIVNAKAGNDDIFV 595
Db 552 EYMTSLIVNGKDTWSV-KGIKNHKGVDYDKLIQFVEKTKHYQARIISLGDKDDVVYS 610
Qy 596 GQGMNIDGGDGHDRVPYSKDGFGNITVDGTSATEAGSYTNARKVARGD--IYHEVVKR 653
Db 611 GAGSSEVFAGEGYDPVSYNKT-DVGKLTIDATGAPKEGYIVPKNM-YGDVEVLQEVYKE 668
Qy 654 QETVKGKRTETIOVRDYELRKVGVOYQSTNLSKVSVEVIGSQFNDVFKSKFNDIFHSG- 712
Db 669 QEVSVGKRTETIQYRDYEFRTGGIPYDVIDNLHSVEELIGKDHDEFKGGKFNDFHGD 728
Qy 713 -----EGDLDLGGAGDDRLFGKGNDRLSGDEGD----- 742
Db 729 GNDYIEGNYGNDRLYGDDGDDYISGGQDDQLFGSGGNDKLSGGGNNYLTGGSGNDELQ 788
Qy 743 -----DLDDGGDDVNLGGAGNDVYIFRKGDNLTLD-GT 778
Db 789 AHGAYNLSGGTGDKLYGGGGIDLLDGGEGNDYLVGGFNDIYVYQNYGHHTIADEGG 848
Qy 779 GNDKLAFADANISDIMITERTKEGIIVKR-----NDHSGSINIPRWYITSNLQNY 827
Db 849 KGDRHLHSDISFDDIAFRVGNLDLIMKAINGVLVSFNESNDVNG-ITFKWFAKD----- 902
Qy 828 QSNKTDHKEIQIKGKGSYITSDQIDKIDKQKGTVTITSQELKKLADENKSQKLSASDI 887
Db 903 ASGADNHLVEITDKGREI---KVDKIPHNNE-----RSGYIKASINASEKNMWNITSV 955
Qy 888 ASSLNKLGVSMALFGTAN 905
Db 956 ANDINKIISVSGFDSGD 973
RESULT 13
S51672
adenylate cyclase hemolysin - Bordetella bronchiseptica
C:Species: Bordetella bronchiseptica
C:Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 16-Jul-1999
C:Accession: S51672
R:Betsou, F.; Sismeiro, O.; Danchin, A.; Guiso, N.
submitted to the EMBL Data Library, September 1994
A:Description: The adenylate cyclase-hemolysin gene from Bordetella bronchiseptica.
A:Reference number: S51672
A:Accession: S51672
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1705 <BET>
A:Cross-references: EMBL:Z37112
C:Superfamily: cyclolysin; calmodulin-sensitive adenylate cyclase catalytic domain homol
C:Keywords: tandem repeat
F:15-328/Domain: calmodulin-sensitive adenylate cyclase catalytic domain homology <ADB>
F:543-1084/Domain: hemolysin A homology <HLYA>
Query Match 17.3%; Score 805.5; DB 2; Length 1705;
Best Local Similarity 27.7%; Pred. No. 2.4e-33;
Matches 254; Conservative 137; Mismatches 334; Indels 193; Gaps 28;
Qy 106 GLDSVENIDRKLKGSANVLSTLSSFLGTALAGIELDSLJ-----KKGDAAP----- 151
Db 403 GYDSLQGV---GSRFSLSGEVSDMAAVEAAEEMTQVHLHAGARQDDAEFGVSGASAHW 458
Qy 152 -----DALAKASIDLINELIIGNLSQ-----STQT-IEAFSSQLAKLG-----STI 190
Db 459 QORALQGAQAAQAQRLVHALMTQFGRAGSTNTPQEAASLSAAVFLGSEASSAVAEIV 518
Qy 191 S-----QAKGFSNTGNKLNQNFNFKTNLGLIEITGLLSGTSAGFALADKNASTGKKV 242
Db 519 SGFFRGSSRWAGGFGVAG-----GAMALGGGIGAVGAGMSLTD-DAPAGOKA 564

| | |
|---|--|
| Matches 231; Conservative 167; Mismatches 355; Indels 302; Gaps 48; | |
| Qy | 43 DFTKAAD-----ELGIARLAEEPHNHTETAKSVDTVNOFL-----SLTQTGIAISATKLEKF 94 |
| Db | 478 DNTKLADGSPAKGYAALAEALDSNGDNIINAADAAFTQLRVWDLNQDNGIS-----528 |
| Qy | 95 LQGHSTNKLAK-GLDSVE-----NIDKLGKASNVLSLSTSLFLGTALAGIELDSLIKGD-148 |
| Db | 529 -QANELRTLEELGIQSLDLAYKDVKNKLG-NWTLAQOQSYTKDGTGTAKMGDLALLAADN 586 |
| Qy | 149 -----AAPDALAK-ASIDLINEI-----IGNLSQSTQTIEAFSSQLAKL 186 |
| Db | 587 LHSRFKDKVELTAEQAANLAGIGRLRDLREAAALSGDLANMLKAYSAAETKEAQLALL 646 |
| Qy | 187 GSTISO-AKFSNIGNKLG--NUNFSKT-NLGLIITGLLSGI--SAGFALADK-----234 |
| Db | 647 DNLHKAETDSDNWKGPSMRLSDTMTQTANEGIALTPSQVQALKGVALVSLSDKAKAAI 706 |
| Qy | 235 -----NASTGKKVAAAGFELSNOVIGNVTKAISYV--LAORVAAAGLSTTGAAVAL 282 |
| Db | 707 DAARDRIAVLDAYTGQDSSTLYMSEEDALNIVKVTNDTYDLAKNIYQNL-----L 758 |
| Qy | 283 ITSSIMLAISPLAFMNAAD-----KFNHA-----NALDEFAKQFR-----KFGY 321 |
| Db | 759 FQTRLQPLNQISFKMENDTFTLDFSLGVQAFNHVKETNPQKAPVDLAEMLAYGELRSWY 818 |
| Qy | 322 DGHLLAEY---QRGVGTIE---ASLTITSTALGAVSAGSAAVGSAGVTPALLVA--373 |
| Db | 819 EGRRLMADYVEEAKKAGKFDYQKVLGQETVALLAKTSGTQADLILQNVGFHNKNSLY 878 |
| Qy | 374 --GVTGLSIGILEASKQAMFESVANRLOGKILEWEKQNGQNY-PDKG-----YDSRY 423 |
| Db | 879 GNDGNDTLIG--AGNDYL-----EGSGSDTYVFKGFGQDTVYNDY 920 |
| Qy | 424 AAYLANNLKFSELNKL-----EAERVI-----AITQORMNNG-----459 |
| Db | 921 ATGRKDIIRFTDGTITADMLTFTREGNHLIIKAKDDSGQVTQVSFYQNDGSGAYRDIHF 980 |
| Qy | 460 -----ELAGITKLGERIKSGKAYADAFEDGKKVEAG-----SNITLAKTGIIID 503 |
| Db | 981 DNGKVLVDVATVKELVQOSTDGSRLYAYQSGNTLNGGLGDDYLYGADGDDLLANGDAGNS 1040 |
| Qy | 504 ISNSNGKKTQALLHPTSPLLTAGTESRERLTNGKYSYINKLFGKRVKQVMTDGEASSKLD 563 |
| Db | 1041 IYSGNGNDT-----LNGEGNDAL-----YGVNGNDALNGG-----1071 |
| Qy | 564 FSKVIQVRAETGDEIGLVNAKAGNDIIFVGQGMNIDGGDHRVFSKDGFGNIT 623 |
| Db | 1072 -----EGNDHL-----NGEDGNDTLIGGAGNDYLEGGSGSDTYVFGK--GFGQDT 1114 |
| Qy | 624 V-----DGTSA-----TEAGSY-TVNRKVARGDIYHEVVKRQETVKGR 661 |
| Db | 1115 VYNYDYATGRKDIIRFTDGTITADMLTFTREGNHLIIKAKDGSQVTVQSYFQNDGSGAYR 1174 |
| Qy | 662 TETIQRYDYELRKVG-----YGVQSTDNLK--SVEEIVGSPQNDVFKG 702 |
| Db | 1175 IDELHFNGKVLVDVATVKELVQOSTDGSRLYAYQSGNTLNGGLGDDYLYGADGDDLLNG 1234 |
| Qy | 703 SKFNDIIFHSGEGDDLLDGGAGDDRLF-----GGKGNDRLSGDEGDDLLDGGSGDDV 753 |
| Db | 1235 DAGNDSIYSGNGNDTLGGEGNDALYGVNGNDALNGEGNDHNGEDNDTLIGGAGNDY 1294 |
| Qy | 754 LINGAGNDVYIFRGDNDTLID---GTG-NDKLAPADANISDIEMIETKEG--IIVKRN 807 |
| Db | 1295 LEGSGSDTYVFGKFGQDTVYNYDYATGRKDIIRFTDGTITAD-MLTFTREGNHLIIKAK 1353 |
| Qy | 808 DHSGSINIPRWYITSNLQNYQSNKTDHKIEOLICKGSIYTSDOOIKLQDKDGTVITS 867 |
| Db | 1354 DDSGQVTQSYF-----QNDGSGA--YRIDIEHFDNGKVLVDVATVKELVQOSTD-----1401 |
| Qy | 868 QELKKLADENKSKLASDITASSINKLVGSMALFG 902 |
| Db | 1402 -----SDRLYAYQSGSTLNGGLGDDYLYG 1425 |

RESULT 17

C81182

iron-regulated protein FrpA, probable NMB0585 [imported] - Neisseria meningitidis (strain
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: C81182
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masinghani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: C81182
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1302 <TET>
A;Cross-references: GB:AE002414; GB:AE002098; MID:g7225809; PIDN:AAF41013.1; PID:g722581/
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0585

Query Match 9.1%; Score 421; DB 2; Length 1302;
Best Local Similarity 22.5%; Pred. No. 9.6e-14;
Matches 198; Conservative 136; Mismatches 292; Indels 254; Gaps 39;

| | |
|----|---|
| Qy | 193 AKGFSNIGNKQLNLFNFKTNLGLIITGLLSGISAGFALADKVAATGKVKVAGFELSNOV 252 |
| Db | 303 AKGFS-----GSLFDHTNNGIRTATGWVSA-DDGLLVRLDN-----GNGI 341 |
| Qy | 253 IGNTVKAISSYVLAQVAAAGLSTTGAVAAALITSSIMLAISPLAFMNAADKFNHANALDEF 312 |
| Db | 342 IDNGAELFGDNT---KLADGSFAKHGYAALAEID-----SNGDNIINAADAAAFQS 389 |
| Qy | 313 AKOFRKFGVDG----DHLLAEYQRGVGTIEASLTITSTALGAVSAGSAAVGSAGVTP 368 |
| Db | 389 LRVWQDLNQDGIISOANELRTLEELGIQSLDLAYKDVKNKLG---NGNTLAQOQSYTKTD- 444 |
| Qy | 369 ALLVAGVTGLISGILEASKQAMFESVANRLOGKILEWEKQNGQNYFDKG--YDSRYAAY 426 |
| Db | 445 -----GTTAKMGDLALLA-----DNLHSRFDKVELTAEQAANLAGIGRLRDLREAAA 494 |
| Qy | 427 LANNLKFSELNKELEA-ERVIAITQ---QRW---DNNIGELA-----GIT 465 |
| Db | 495 LSGDLANMLKAYSAAETKEAQLALLNLHKAETDSNMWKKSPMRLSTDTWTQTANEGIA 554 |
| Qy | 466 KLGERI-----KSGKAYADAFEDGKKVEAGSNITLDAKTG-----500 |
| Db | 555 LTPSQVAQLKKNALVSLSDKAKAAIDAARDIAV-----LDAYTQDSNTLYMSEED 607 |
| Qy | 501 ---IIDISNSN-----GKKTQALHFTSPLLTAGTES 528 |
| Db | 608 ALNIVKVTNDTYDLAKNIYQNLFFQRLQPLYNQISFKMENDTFTLDF-SGLVQAFNHV 666 |
| Qy | 529 RELTNGKYSYN---KLKFGRYKNW-----QVTDGEASSKL-DFSQVJOR-----570 |
| Db | 667 KE--TNPQKAFVDLAEMLAYGELRSWYEGRLMTDYVEEAKKAGKFDYQKVLGQETVAL 724 |
| Qy | 571 VAETEGTDEIGLVN-----AKAGNDIIFVGQGMNIDGGDHRVFSYKDG 618 |
| Db | 725 LAKTSGTQADDIILQNVFGHNKNVSLYNGDNDTLIGGAGNDYLEGGSGSDTYVFG--G 782 |
| Qy | 619 FGNITV-----DGTSA-----TEAGSY-TVNRKVARGDIYHEVVKRQET 656 |
| Db | 783 FGQDTVYNYDYATGRKDIIRFTDGTITADMLTFTREGNHLIIKAKDGSQVTVQSYFQNDG 842 |
| Qy | 657 KVGRKRTIQRDYELRKVG-----YGVQSTDNLK---SVEEIVGSGFN 697 |
| Db | 843 SGAYRIDEIHFDNGKVLVDVATVKELVQOSTDGSRLYAYQSGNTLNGGLGDDYLYGADG 902 |
| Qy | 698 DVFKGSKFNDI FHSGEGLDLDGAGDDRLF-----GGKGNDRLSGDEGDDLDG 748 |

Db 903 DLNGDAGNDISYSGNGNDTLDDGEGNDALGYNGNDALNGEGNDHNGEDGNDTLGG 962
Qy 749 SGDDVNLGAGNDVYIFRKGNDNTLYD---GTG-NKLAPADANISDIMIERTKEG--I 802
Db 963 AGNDYLEGGSGDYVFGKFGQDVAVNYDVTAKDIIIRDTGITAD-MLTFTREGNHL 1021
Qy 803 IVKENDHSGSINIPRWYITSLNQVSNKTHKIEQLGKDGYSITSDOIDKILQDKKG 862
Db 1022 LKAKDGGQVTVOSYF-----QNDGSGA---YRIDEIHFNDKGVLDVATVKELVQOSTDG 1074
Qy 863 TWITSQELKLADENKSKLSASDIASSLNKLGVSMALFG 902
Db 1075 -----SDRLVAYSGNTLNGGLGDDVLYG 1098

RESULT 18
B82736
hemolysin-type calcium binding protein XF1011 [imported] - Xylella fastidiosa (strain 9a
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: B82736
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: B82736
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1636 <SIM>
A:Cross-references: GB:AE003938; GB:AE003849; NID:g9105935; PIDN:AAF83821.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1011

Query Match 8.9%; Score 415; DB 2; Length 1636;
Best Local Similarity 22.4%; Pred. No. 2.7e-13;
Matches 226; Conservative 150; Mismatches 323; Indels 308; Gaps 48;
Qy 25 LYLAIKPDYDQKGTLDNFIKADELG---IARLAEPNITETAKKSVDTVNFQSLTQ 81
Db 793 MYLALP-----PQLSATFAFAEAVAKKESYIRLVLEPRUTDYLSC-----LRUTD 841
Qy 82 TG--IAISATKLEFLQKHSNTKLAKGLDSVENIDRKLKGSANVLSTLSSFLGTALAGIE 139
Db 842 NNGVMWADASGLEAKLDQTHQNKAAQLQDVMWDLR---YGSNAV-----AASGWK 889
Qy 140 ----LDSLIKGDAAPD---ALAKASIDLNEITIGNLSQSTQTTAEFSSQLAKLGSTISQ 192
Db 890 PFDALRHMDIRTAATPDGRQALAAAEITLVS---GNAEGS-----DAADLLFGDAGANLLR 942
Qy 193 AKGFSNIGKNLQNLNFSKTNLGLIITG-----LLSGISAGFALADKNASGKVAAGFE 247
Db 943 GGG-----GDDVLSCGGGNDTLGEGAGNDTL---YGDGDDVDLGGE 981
Qy 248 LSNQVIGNVTXAI--SSVYLQORVAAGLSTTGVAALITSSIMLAISPLAFMNAADKPNH 305
Db 982 GSNRLEGGAGNDVLKVSWSADNVILG---GTGDDTLGYS---AFADTYLFGQGD--GH 1032

Qy 306 ANALDEPAKQFRKFG-----YDGHLLAEYQRGVGTI-----EAS 340
Db 1033 DTIIEOGQTKLVFEGIGLAADVRLIREGQDVLDLNGHDSIRLKDWLTSNGTRNHAD 1092
Qy 341 LTTISTALGAVSAGVSAAVG-SAVGTPIALLVAGVTG---LISG-----ILEASKQAMFE 392
Db 1093 IEQIVFADGTLTWTPETLSSMGLTTLGTPGNDTLKGWQKDIILGGAGDDVLDGEG--- 1148
Qy 393 SVANRLOG-----KILEWEKON---GGQNYFDKGYDSRYAAYLANNLKFSELNKELE 442
Db 1149 --SNRLEGGAGDDVLKVSWSADNVLSGG-----TGDDTLGSAFADTYLF----- 1192
Qy 443 AERVIAITQORWNNNIGELAGITKL--GERIKSKAYADAFEDGKKVACASNITLDAKTG 500
Db 1193 -----NQGDGHTTIEOGGTDKLVFEGIGLAA-----DVRLLREGQDVVLD--- 1233
Qy 501 IIDISNSNGKKTQALHFTSPLLTAGTESRRLNGKYSINKLKFGRVKNQVQVTDGSSAS 560
Db 1234 -----LNGHDSI-----RLKDWLTSNGTRNH 1255
Qy 561 KLDFSKVIRVAETEGT-----DEGLIVNAKAGND-----DIFV-GQGKMMIDG 604
Db 1256 SAD-----IEQIVFADGTLTWTPETLSSMGLTTLGTSNDTLKGWQKDIILGGAGDDVLDG 1311
Qy 605 GDGHDV-----FYSKD---CGFGNITVDGTSATEAGSYTVNRKVARGDIYH 648
Db 1312 GEGSNRLEGGAGDDVLKVSWSADNVILGTDGDTLYGSAP--ADTYLFN-----KGHDH 1365
Qy 649 EVVKRQET-----KVGKRTETIQVRYELRKYGVGYOSTD 683
Db 1366 TIIEOGQTKLVFEGIGLAADVRLIREGQDVLDLNGHDSIRLKD--LTSNGTRNHAD 1424
Qy 684 NLKSVEEVI-----GSQFNDVFGSKFNDI--FHSGEGDDLLDGCAG 723
Db 1425 -----IEQIVFADGTLTWTPETLSSMGLTTLGTPGNDTLKGWQKDIILGGAGDDVLDG 1480
Qy 724 DDLRFGKKGND--RLSGDEGDDLLDGGSGDVLNGGAGNDVYIFRKGDGNDTLVDGTGND 781
Db 1481 SNRLEGGAGDDVLKVSWSADNVLSGGTGDYTLGSAFADTYLFGKDGHTTIEOQGTG 1540
Qy 782 KLAF-ADANISDIMIERTKEGIIIVKRNHSGSINIPRWYITSLNQVSNKTHKIFOLI 840
Db 1541 KLVFAGLSRKEARFTSGDDLSLFLNGSEDQVTVACWF-----NGSGHGVESLV 1590
Qy 841 GKGSYITSDQIKLQDKDGTGVTISQELKLADENKSKLSASDI 887
Db 1591 FQDGT-VLSAEVERLIAAMALSPAVITMQ-ASVGETKESPRLVASSI 1635

RESULT 19

C82779
hemolysin-type calcium binding protein XF0668 [imported] - Xylella fastidiosa (strain 9a
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: C82779
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: C82779
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1208 <SIM>
A:Cross-references: GB:AE003910; GB:AE003849; NID:g9105532; PIDN:AAF83478.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr

Db 544 SNWKKSPRLSTDTWTANEGIALTPSOVAQLKGNALVSLSDKAKAIDAARDRIAV-- 601
Qy 490 GSNITLDAKTG-----TIDISNSN----- 508
Db 602 -----LDAYTQDSSTLYMSEEDALNIVKVTNDTYDHLAKNIYNQLLFQTRLQPLNQI 656
Qy 509 -----GKKTQALHFTSPILTATGESPERLTNGKYSIN---KLKGRVKNW-----Q 552
Db 657 SFKMEKNDFTLDF--SGLVQAFNHVKE--TNPQAFVDLAEMLAYGELSRWYEGRRLMADY 713
Qy 553 VTDGASSKL--DFSVKIOR-----VAETEGTDEIGLVN-----AKAGNDDIF 594
Db 714 VEEAKKAGKFDYQKVLQOETVALLAKTSGTQADDILQNVGFHNKVNLSVYNGDNDTLI 773
Qy 595 VQGGKKNIDGGHGRVRYSKGGGNTIV-----DQTS-----TEAG 633
Db 774 CGAGNDYLEGGSGSDTYVFGK--GFGQDTVYNYDYATGRKDIIRPTDGITADMLTFTREG 831
Qy 634 SY-TVNRKVARGDIYHEVVKQETVKGRKETIQRDYELRVKG----- 676
Db 832 NHLTIKAKDSQGVTVQSYFQNDGSGAYRIDEIHPDNGKVLVDVATVKELVQOSTDGSRL 891
Qy 677 XYQYSTDLNK---SVEEVIGSFQNDVFKSGKFNDFHSGEGDDLLDGGAGDRLF----- 728
Db 892 YAYQSGSTLGLGDDYLYGADNDLLNGDAGNDSIYSGNGNDTLDGEGNDALYNGN 951
Qy 729 ---CGKNDRLSGDEGDDLLDGGGDDVNLGGAGNDVYIFRKGDNDTLVD---GTGND 781
Db 952 DALNGEGNDHLNGDNDTLGGAGNDYLEGSGSDTYVFGEGQDTVYNYHYVDKNSD 1011
Qy 782 KLAADANISDIMIERTKEGIIIVKENDHSGSINIPRWITSLNQYQSKTDHKEQLIG 841
Db 1012 TMHFKGFAADVHFIRSGSDLVLSAQD--NVRISGFFYGEN-----HRVDTFV 1060
Qy 842 KDGSYTSDQDKILQDKKQGTVITSQELKXLADENKSKLSASDIASGL---NKLVGSM 898
Db 1061 DDA-----TANSSSNALQPIPTOQTGILAPS 926
Qy 899 ALFG-----TANSSSNALQPIPTOQTGILAPS 926
Db 1085 SVFGSNTAATGNGVDAN--IQSVQQLP--LVTPS 1114

RESULT 21

G82562
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: G82562
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: G82562
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-2064 <SIM>
A:Cross-references: GB:AE004049; GB:AE003849; NID:g9107579; PIDN:AAF85206.1; GSPDB:GN001
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tauhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2407

Query Match 8.4%; Score 391; DB 2; Length 2064;
Best Local Similarity 25.4%; Pred. No. 6.5e-12;
Matches 135; Conservative 89; Mismatches 163; Indels 144; Gaps 23;
Qy 446 VIAIQRWNNIGELAGITKIGRIKSGKAYADAF-----EDGKK--VEAGSNITLDAKT 499
Db 1588 VLKVAVSADNVL-----IGTGDDTLGSAVADTYLNFNKGDDHTTIEQGGDTLVFGA 1642
Qy 500 GIIDISNSNGKQALHFTSPLLTAGTESRERTNGKYSYNKLKFRGVKNWQVTDGEAS 559
Db 1643 GIV-----ASQV-----RVLRGQDVVLDLNGHDSI-----RLKDLTSDGYRN 1682
Qy 560 SKLDFSKVIORVAETECT-----DEIGLIVNAKAGNDDIFVQGGKRNIDGGDGHDRVF 612
Db 1683 GNND---IEQIVPADGTIWTPELSSMGLTTLTGSGNDTLKGWQKDIILGGDGDVIL- 1737
Qy 613 YSKDGGFG-----NITVDGT-----SATEAGSYTVNRKVARGD 645
Db 1738 ---DGGMSNRLEGGAGNDVLKVAYSADNVILGGTGDDTLGSGYADTYLNFN-----KGD 1790
Qy 646 IYHEVVKR--QETKV-----GKRTETIOYRDYELRVKVGYYQ 680
Db 1791 GHDTLIEQGGDDTLVFGAGIVASQVRVREGQDVLDLNGHDSIRLKW---LTSQYR 1847
Qy 681 STDNLKSVEEVI-----GSQFNDVFKSGKFNDFHSGEGDDLLDG 720
Db 1848 NGNN--DIEQIVPADGTIWTPELSSMGLTTLTGSGNDTLKGWQKDIILGGDGDVLDG 1905
Qy 721 GAGDRLFGKGN--RLSGDEGDDLLDGGGDDVNLGGAGNDVYIFRKGDNDTLVDGT 778
Db 1906 GMSNRLEGGAGNDVLKVAYSADNVILGGTGDDTLGSAVADTYLNFNKGDDHTTIEQ 1965
Qy 779 GNDKLAADANISDIMIERTKEG--IIVKRNDSHSGSINIPRWITSLNQYQSKTDHKE 836
Db 1966 GDDTLVFG--AGLHQKEARFTKSGNDLSILFNASEDQVTIAGWF-----NGSGHOV 2014
Qy 837 EQLIGKDGSYTSDQDKILQDKKQGTVITSQELKXLADENKSKLSASDI 887
Db 2015 ESLVFDQGT-VLSGEVERLIAAMALSSAVITQ-ASVRDTKESHRLVASSI 2063

RESULT 22

S34238
leukotoxin A - Pasteurella haemolytica (fragment)
N:Alternate names: lktA protein
C:Species: Pasteurella haemolytica
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Mar-2001
C:Accession: S34238; S34236
R:Lainson, A.F.; Aitchison, K.D.; Donachie, W.
submitted to the EMBL Data Library, June 1993
A:Description: DNA sequence of the carboxy terminal end of leukotoxin A from the T4 sero
of Pasteurella haemolytica.
A:Reference number: S34236
A:Accession: S34238
A:Molecule type: DNA
A:Residues: 1-208 <LA1>
A:Cross-references: EMBL:Z22885; NID:g311830; PIDN:CAA80499.1; PID:g311831
A:Experimental source: serotype T4
A:Accession: S34236
A:Molecule type: DNA
A:Residues: 9-208 <LA2>
A:Cross-references: EMBL:Z22886; NID:g311826; PIDN:CAA80500.1; PID:g311827
A:Experimental source: serotype T15
C:Genetics:
A:Gene: lktA
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat

Query Match 8.2%; Score 383; DB 2; Length 208;
Best Local Similarity 37.8%; Pred. No. 6.le-13;
Matches 79; Conservative 59; Mismatches 64; Indels 14; Gaps 8;

Qy 723 GDBRLFGKGNDRLSGDGDDLLDGGSGDDVYANGCAGNDVVIFRKGDNNDTLVDGTGNDK 782
|:|||||||:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 1 GNDRLFGKGGDDIYDGGNGDDPFDGGKGNOLLHGKGGDDIFVHRQGNGDSITESGNDK 60
|:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Qy 783 LAPADANISDIMITERTKEGIIVKRNDHSGSINIPRYW---ITSNLQNYOSNKTDKHIEQ 838
|:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 61 LSFSDSNLKDLTEKNVHLVI-TNTKQEKVTQNWFREAEFAKTIYNVATR-DKLEE 118
|:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Qy 839 LIKGGSYIITSQQIKLOPKDGTVTTSQELKKLADENKSQKLSDIASLNKLVGSM 898
|:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 119 IIGQGERITSKVDELIEKGK--KIDKSDLQWVDNYQLLYVS-RDASNSLDKLISA 175
|:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Qy 899 ALPGTANSVSSNALQPITQTPTQGIAPSV 927
|:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 176 SAFTSSND-SRNVL---ASPT-SMLDPSL 199

RESULT 23
F81856
probable RTX-family exoprotein NMA1626 [imported] - Neisseria meningitidis (strain Z2491)
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: F81856
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
C.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:2022556; PMID:10761919
A:Accession: F81856
A>Status: preliminary
A:Molecule type: DNA
A:Amino acid type: P
A:Residues: 1-697 <PAR>
A:Cross-references: CB:AL162756; GB:ALL57959; NID:g7380091; PIDN:CB84854.1; PID:g738026
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1626

Qy 455 DNNIGELAGTIGKERIKSKAVADAFEDCKKVEAGSNITLDAKTGIIDISNNGKKTQA 514
Db 294 -----DYEEAAKAGK-----PDKYKVLGETVALLAK----- 322
Qy 515 LHFTSPLLTAGTRESRELTNGKYSINKLAFGRVKWQVTDGEASSKLDFFSKVIQVRAET 574
Db 323 -----TSGTQADDILQN-----VGFHNKN-----VSL 345
Qy 575 EGTDEIGLIYNKAGANDDIIFVGQGNIDGGDGHDRVIFYSKDGGFGNIV----- 624
Db 346 YGND-----GNDTLIGGAGNDYLEGSGSDTYVFGK--GFGQDVTYNYDYATGRK 393
Qy 625 -----DGTSA-----TEAGSY-TVRKVKARGDIYHEWVKRQETKVKGRFTETIQRYDEL 672
Db 394 DIIRFTDGIATDMLTFTREGNHLILKAKODSGQVTVQSYFQNDGSGAYRIDETIHFDMGKV 455
Qy 673 RKVG-----YGVQSTDNLK--SVERIVGSQFNDFPKSGKFNDFPHSGE 713
Db 454 LDVATVKELVQOSTDGSRLIYAVQSGNTLNGGLGDDYLYGADGDDLLNGDAGNDSIYSGN 513
Qy 714 GDDLLDGGAGDRLF-----GGKGNDRLSGDEGDDLLDGGSGDDVLNGAGNDVYI 764
Db 514 GNDTLGEGNDALYGYNGNDALGGEGNDLNGEDGNDTLIGGAGNDYLEGSGSDTYV 573
Qy 765 FRKGDGNDTLYD---GTGNDKLAFAADANISDIEMIERTKEGIIIVKNDHSGSINIPRWIIT 821
Db 574 FGKFGQDQVTYNYVHVDKNSTDMFIRKFGKAADVHFIRSGDVLVLSASEQD-NVRISGFYFG 632
Qy 822 SNLQNYQSNKTDHKIEQLKOGSYYITSDQIDKILQDKDGTVITSQELKKLADENKSQK 881
Db 633 EN-----HRVDTFFVDDAA----- 646
Qy 882 LSASDIASSL---NKLVGSMALFG-----TANSVSSNALQPIITQPTQGIILAPS 926
Db 647 ISNPDPFAKYNAGNNLVQSMVSGSNTAATGNGVDAN-IQSVQQL--LVTPS 696
RESULT 24
AH2515
hypothetical protein alr7304 [imported] - Nostoc sp. (strain PCC 7120) plasmid
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AH2515
F:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Yamamoto, S.; Watanabe, A.;
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.;
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobact
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2515
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-4936 <KUR>
A:Cross-references: GB:BA000020; PIDN:BAB78388.1; PID:gl7135842; GSPDB:GN00180
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr7304
A:Genome: plasmid

16

Db 3781 AIDYTTTADQASPSQRHPELDDIFAPGILITGANANGG-----TTTLGGTSQAT-----AY 3832

Qy 190 SSQAKLGGSTISQAKFSNIGNKL-----QNLNFSKTNLG----- 214

Db 3833 LTGVATLAAQIAQEK-----LGRKLTVTFRNLLDTTSVIINDGNDENVNTNTGTFNPRV 3888

Qy 215 --LEIITGLS-----GISAGFALADKNASTGK-----KVAAGFELSN 250

Db 3889 DLLKLAELSLTGTTPNDPVPNG-----NNNNNGTTTSNTINQVHTVNLAAQVQRTD 3944

Qy 251 QVIGNVTKAISVYLAQRAAGLSTTGAVAAALITSSIMLAISPLAFMNAADKFNHANALD 310

Db 3945 VDFGN-----QQIITNQAPT--VANAIADQIINEDANFTFVIPANTFVDADAGD 3991

Qy 311 -----EPKQRPKF-GYDGDHLLAEYQGVGTIEASITTTISTALGAVSAGVSA 357

Db 3992 VLTYSTTLPSWLTFNATRTFSGPGN-----SNVGTVAITVT-----AT 4031

Qy 358 AAVGSAGCTPALLVAGVTGL-ISGILEASKQA-----MPESVANRLQKILEWEKQNG 410

Db 4032 DSTGASVDDSTFLTVANTNDAPILGLAIADQSTASNTFTFQIPLNTFSD----- 4081

Qy 411 GONYFDKGYDKRYAAYLANNLKFLSELNKELEAERVIAITQORWNNNIGELAGITKLGER 470

Db 4082 ----IDTGDTLTYSKLVGDIPLPTWL-----TFNATNRTFSGIPG----- 4118

Qy 471 IKSGKAYADAFEDGKKVBAQS-NITLDAKTGIIIDISNGSKKTQALHTSPL----- 521

Db 4119 -----NVDVGTNLNTVQA-----IDTSNASISDSFVLTTITNLINNIVGTSG 4159

Qy 522 --LTAGTESREL--TNKYSYINKLKEGRVKNQVTDGEASKL-----DF 564

Db 4160 NNTLGTTPNNNOIGLGG-----NDIIFLAGNDTLNGGTGSDTMTWGLGDDTVIVDNNV 4214

Qy 565 SKVIOVAETEGTDEI-----GLI-----VN 585

Db 4215 DKVVENL--NEGIDTVRSISVTLLENVENLILGTSTNIGTILNLSIITGNSCANTLN 4272

Qy 586 AKAGNDDIFVQG-GKONTIDGGDHDRVYKDGFGNITVDTGTSATEAGSVTVNRKVARG 644

Db 4273 GKAG-DDILNEGNDNLKGBDGNVL-----NGGAGNDILGGLGDDVMTGGVG-----N 4322

Qy 645 DIYH-----EVRKQETKVKRTETIOYRDYELRVKG-YGYQSTDN-LKSVBEVI 692

Db 4323 DIYYVDSNDIIDEINSGTDTVTNTIITWTLGNHLENLTLTGSSAINGTGNALKNI--II 4380

Qy 693 GSQFNDVFKSKFNDIPIHSGGDDLLDGGAGDRLFGKGNDRLSGDEGDDLLDGGSGDD 752

Db 4381 GNSADNLSGGDNDILRGGEGNDTLGGAGNDSLGGIGNDSLNGEDGNDNLKGDVGN 4440

Qy 753 VLNGAGNDVYIFRKGDNNDTLYGTDGNDKLAPADANISDMIRTEKGIIVKRNDSHGS 812

Db 4441 ILNGNAGNDT--LDGGLGDVMTGAGND-IYFVDS--NDTIIBELNEG-----TDTVNAS 4492

Qy 813 INIPRWYITSLNQNYSKNDHKIEQLIGKGSYITSQIDKILQDKDGTIVITSQELKK 872

Db 4493 IN--WTLGNLENL-----TLTGSGINGTGNALKNI----- 4522

Qy 873 LADENKSKLASDIFASSINKLVGSMALFGTANSVS 908

Db 4523 ITGNNGDNLGGDNDTLRGNAGNDTLFGSGGNS 4558

RESULT 25

AH3098

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C:Accession: AH3098

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.; erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyayavin, T.; Levy, R.; Li, M.; McClellan; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AH3098

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1944 <KUR>

A:Cross-references: GB:AE008689; PIDN:AAL45206.1; PID:gl77742885; GSPDB:GN00187

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: rzCA

A:Map position: linear chromosome

Query Match 7.1%; Score 330.5; DB 2; Length 1944;

Best Local Similarity 21.7%; Pred. No. 7.7e-09;

Matches 218; Conservative 147; Mismatches 333; Indels 305; Gaps 48;

Qy 66 AKKSVDTVNQFLSLTQGT-----IAISATK-----LEKFLQKHST-NKLAKGLDSVENI 113

Db 653 ARWADAGDDIVASAAAGSAGVANVALSAKLSQDITRSFWTDSNGDVVAEGKEAL--A 710

Qy 114 DRKLKASNVLSLSSFLGTALAGIE-----LDSLIKKGDAAPDALAKASIDLINEIIGNL 169

Db 711 DLSLGKHKLTLT-----TDVNGVKTSDTLDLVL-RGD----- 742

Qy 170 QSQTQTEAFSSQAKLGSTISQAKGF-----SNIGN-KLQNLNFSKTNLGLIETGLLSG 224

Db 743 --RTLLVDNFND-----GKADGWATDFGATDGTGNFLKGTGFVSRTTG-----DG 787

Qy 225 ISA-GFALADKNASTGKVAAGFELSNOVLGNVTKATISSYVLAQRAAGLSTTGAVAAI 283

Db 788 LDPEAALFQDSRATGNKLV-----YIGAOSGWSNYVF-----EATL 825

Qy 284 TSSIMLAISPLAFMNAADKFNHANALDEFKQFRKFGYDGD-----HLLAEYQGVGTIEAS 340

Db 826 TOLDNDAIG--VYFYSD-----AKNYFRFTMDGEANRRRLVKVNSGVETILLAS 872

Qy 341 LT-----TISTALGAVSAGVSAAGVAVGSAVGTPIALLVAGVTGLISGILEASKQA 389

Db 873 VNEGSPYNDIPLTVATVDGAINVFLGDKN-----FGGPVVDVTSPLSGGTGVLGYSSQRA 929

Qy 390 -MPESVANRLQKILEWKQNGQYEDKGVDSRYAAYLANNLKFLSELNKELEAERVIA 448

Db 930 SVFDDI--VVTKAATTAKAGIDORVYDFGDKG-----VSVRLDASSSFS 972

Qy 449 ITQOR--WDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSN----- 492

Db 973 AGQLRDPFWTDLGNNVVAI-----GKKVDADNLVGVNKLMLMVSDA 1013

Qy 493 -----ITLDAKTGIIDISN-----SNGKKTQALHFTSP 520

Db 1014 QGSVSTDRIDVTVVVDKTKILVAENFSTAEAMARFKIVDEGEFGGIGADGKSESLISDGK 1073

Qy 521 LL-TAGTESREL--NG-----KYSY-----INLKFGRVKNQVMTDGEASSKLDFSK 566

Db 1074 LLQTTGLMSRELTVNGATNDSKTKRGWSPMGDGVNLRGLTYALF--NDPAAQAWTDYA- 1130

Qy 567 VIQVAAETEGTDEIGLIIVNAKAGNDDIFVGGKGNIDGSDGHDRVYFSKGGFNI----- 622

Db 1131 -IEANFQTPDKDGLGFLFRYKDSKN-----YYKLELDADGILDR--NPGNGAGSIFNLV 1181

Qy 623 ----TVDGTSATEAGSVTVNRKVARGDIYHVVVKRQETKVKRTETTYQ--RDYELRVKG 676

Db 1182 RMKNGIBEILAAQVPEKPEQGE-----KLREVEVIGDKITAFINEBALFAYPIDGRGLAAGT 1238

Qy 677 YGYOSTN-----LKS---VEEVIGSFQDNVFKGS-----KFNDIPIHSG 712

Db 1239 FGLYSWGNAGLTFNLTVDLKSGLFVNRIVGTNGADVLTGTAATMTFGLGNDLQGF 1298

Qy 713 EGDLLDGGAGDRLFGKGNDRLS-----GDEGDDLLDGGSGDDVLNGAGNDV- 762

Db 1299 GGDDRLOGGAGDADALAGGTGNDALSGGAGNDELWGDGDDMIAGGLGDDTFEGGRGNDLL 1358

Qy 763 -----YIPKGDGNDTLYD---GTGN-DKLA PADANISDI MIERTKEGIIIVKRNDS 810
Db 1359 IGGDGSRYRYGRGDGSDVIIVETASGSDVRLSLYDIDRSEAVLRKYGVSVVIELADGE 1418
Qy 811 GSINIPRWYITSNLQYOSNKTDKHKBOLIGKGSYITSDQIDKILQDKKDGTVITSQEL 870
Db 1419 -----TLTSLN---QLADGGERLSPADGVWLNRGDIVKGLVNR--GPVAADDGL 1463
Qy 871 KKLADENKSKLS-ASDIASSLNKLVGSMALFGTANSVSSNAL 912
Db 1464 AAVNEDAPSFVISPATLLGNDRDADLDGLTAVTGVSAVPVGGTAV 1506
RESULT 26
A96188
probable phosphatase (EC 3.1.1.-) yvnb [imported] - Agrobacterium tumefaciens (strain
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C;Accession: A96188
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: A96188
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1990 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK89027.1; PID:g15158819; GSPDB:GN00170
C;Genetics:
A;Gene: AGR L 909
A;Map position: linear chromosome

Query Match 7.1%; Score 330.5; DB 2; Length 1990;
Best Local Similarity 21.7%; Pred. No. 8e-09;
Matches 218; Conservative 147; Mismatches 333; Indels 305; Gaps 48;
Qy 66 AKKSVDTVNFQSLTQTQ-----TAISATK-----LEKFLQKST-NKLAKGLDSVENI 113
Db 699 ARAMADAGDDIVASAASGASVANVALSAAKSLRSQDINSFVWTDNSGVDVAEGKEAI--A 756
Qy 114 DRKLGKASNVLTSLSSFLGTALAGIE---LDSLIKKGDAAPDALAKASIDLINEIIGNL 169
Db 757 DLSLQKHKLTIT-----TDVNGVKTSDTLDLV-RGD----- 788
Qy 170 SQSTQTIEAFPSQLAKGSTISQAKGF-----SNIGN-KLQNLNFSKTNLGLIEIITGLSG 224
Db 789 --RTLLVDNFND-----GKADGWATDFGATDGTGNTGNFLKGTVPFSRSTTG-----DG 833
Qy 225 ISA-GPALADKVAAGTGGKVAAGFELSNOVIGNVTKAISSVLAQORVAAGLSTTGAVAAI 283
Db 834 LDAPEALFDQSDATGNKLV-----YIGAQSNGMSNYVP-----EATL 871
Qy 284 TSSIMLAISPLAFMNAADKFENANALDEFAQFRKFGVDGO---HLLAEYQRGVGTIEAS 340
Db 872 TQLONDAIG--VYFYSD-----AKNYRFTWDGAEANRQLVKVNGVETILLAS 918
Qy 341 LT-----TISTALGAVSAGVSAAGVAVGTPIALLVAGVTGLISGILEASKQA 389
Db 919 VNEGSPYNMDIPLTVAIVDGAIVNPLGDKNV---FGGPVVDVTSPLSGGTGLVYSSGQRA 975
Qy 390 -MPSVANRLQKILEWEKQNGQNYFDKGYDSRYAAYLANNLKFLSELNKLAEARVIA 448
Db 976 SVFDDI---VVTKAATTAKAGIDQRYVDFDGDGK-----VSVRLDASSFS 1018
Qy 449 ITQOR---WDNNIGSLAGITKLGRIKSGKAYADAFEDGKKVEAGSN----- 492
Db 1019 AQGLRDFVWTDLGNVVAI-----GKKVDADLNVGVNKLMLMSDA 1059
Qy 493 -----ITLDAKTGIIIDSN-----SNKKTOALHFTSP 520
Db 1060 QGSVSTDRIDVTVDKTKILVAENFSTAEAMARFKIVDEGFGGIGADGKSSWELISDGK 1119

Qy 521 LL-TAGTESRRLTNG-----KYSY-----INKLKFRGVKNQVTDGCEASSKLDPSK 566
Db 1120 LLQTTGLMSELTLWNGATNSDKYRGWSPMGWDGVNVLRTGYALP--NDPAAQAWTDYA- 1176
Qy 567 VIQVABETGTEIGLIVNAKAGNDIFVQGGKRNIDGGDGHDRVYFVKDGGFQNI----- 622
Db 1177 -IEANFQTPDKDGLGFLFRYKDSKN-----YKLELDADGILDR---NPNAGAGSIFNLV 1227
Qy 623 -----TVDGTSATEAGSYTVNRKVARGDIIYHEVVKRQETKVKRTETIQY--RDVELRKVG 676
Db 1228 RMKNGIEEILAQVPGKYEPQEM---KLREVEVIGDKITAFINEEALPAYPIGRGLAAGT 1284
Qy 677 YGYQSTDN-----LKS---VEEVIGSFQNDVFKGS-----KFNDFHSG 712
Db 1285 FGLYSWGNAGLTPDNLTVVLDKSGLEVNIRVINGADVLVGTAAETMFGLEGNDLQGF 1344
Qy 713 EGGDLLDGGAGDRLFGKGNDRLS-----GDEGDDLLDGGSGDDVLNAGAGNDV- 762
Db 1345 GGGDRLDGGAGDALLAGGTGNDALSGGAGNDELWGGDDGDDMIAGGLGDDFIEGGRGNDLL 1404
Qy 763 -----YIFRKGDNNTLYD---GTGN-DKLA PADANISDI MIERTKEGIIIVKRNDS 810
Db 1405 IGGDGSRYRYGRGDGSDVIIVETASGSDVRLSLYDIDRSEAVLRKYGVSVVIELADGE 1464
Qy 811 GSINIPRWYITSNLQYOSNKTDKHKBOLIGKGSYITSDQIDKILQDKKDGTVITSQEL 870
Db 1465 -----TLTSLN---QLADGGERLSPADGVWLNRGDIVKGLVNR--GPVAADDGL 1509
Qy 871 KKLADENKSKLS-ASDIASSLNKLVGSMALFGTANSVSSNAL 912
Db 1510 AAVNEDAPSFVISPATLLGNDRDADLDGLTAVTGVSAVPVGGTAV 1552

RESULT 27
C82521
hemolysin-type calcium binding protein XP2759 [imported] - Xylella fastidiosa (strain 9a
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: C82521
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82521
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1296 <SIM>
A;Cross-references: GB:AE004081; GB:AE003849; NID:g9107985; PIDN:AAF85544.1; GSPDB:GN001
A;Experimental source: strain 9a5C
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorfy, H.; Facinani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
Rodrigues, V.; Rosa, A.J. de M.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2759

Query Match 7.1%; Score 329; DB 2; Length 1296;
Best Local Similarity 21.6%; Pred. No. 5.1e-09;
Matches 215; Conservative 146; Mismatches 373; Indels 260; Gaps 46;
Qy 77 LSLTQTGTGIAISATKLEKFLQKHSNKLAKGLDSVENIDRLKGSANVLSLSSFLGTALA 136

Db 178 LDLDGSGIETGSDGRVLFHDADGVKTGTCGLXPDDGWLVDNRNGTIDS--GRELF 235
QY 137 GIELDSLKKGAAPDAL-----AKASIDLINEIGNL-----SOSTQTI 176
Db 236 G--TDLKRGQLATDGFALRDVDSNQDKIDAADRVPANLRWRDLNQDQISQANELS 293
QY 177 EAFSSQLAKLSTISQAKGFSNIGN-KLQNL--NFSKTNLGLLEIITGLLSISAGF-ALA 232
Db 294 LLDANHHVIG--VTATAGRVDLGNVQTAAGTFRSN-GTTGATEGTTGTAANLDLLA 350
QY 233 D-----KNASTGKKAAGPELSNOVIGNVTKAIS-----SYVLAQRVAAGL- 273
Db 351 DTFYRDFTCQOVALTOAKALPLRSGSRVRLNDEAISLKDGNVWQSYSQOSTFOAQOLD 410
QY 274 -----STTGAVAALITSSIMLAISPLAFPMNAADKFNHANALDE-----FAKQFR 317
Db 411 RLDGLMEKWANTSDMSLRAQAEALASKGVTVYVLSGLNPGTAAYQDFLRKLGIVERFM 470
QY 318 KFGYDGDHLLAAYQYGVGTIEASLTITISTALGAVSAGVSAAGVSGVGTPIALLVAGV-T 376
Db 471 GFTYGG-----QQG-----EARFPLDATSGTMTVSLSDVQV-----TNIALAYERPKT 514
QY 377 GLISGILEASQAMPESVANR--LOGK-ILEMEKONGQNYFPDKGYDSRYAAYLANNLKF 433
Db 515 DYESLLKTRMEPIYNLAEDFVNGAWMDW---SGVERALKQGIQRHPRDGIILDAIEF 571
QY 434 LSELNKELEAER-----VIAITQORW----- 454
Db 572 VSALGYK-TAERLGNWNAIFGLADQLSAAPDMGAFDHSSSWTVIPAAADRHFITGSARSD 630
QY 455 -----DNNIGELAG---ITKLERIKSKAYADAFEDCK-----K 486
Db 631 VLLGTSGNDMIIGEVTVGNDVLIGKGNNDTLOGSGGDDTLDGCTGNDTLNGAGNDTYRFA 690
QY 487 VEAG-----SNITLDAKTGIIDISNSGKKTQALHFTSPLLTAGTESRERTNGKYSY-- 539
Db 691 ICAGVDSIYEDVTATSDTIVFADVRSTALTALERKDYDLVIKYGTSQDLTINNYFPC 750
QY 540 -----INKLKFRGRVNWQ-----VTDGEASSKL--DFSKVIOQVVAETEGTDEI--GL 582
Db 751 YSGAKIEQTFSDGVTWDDAAIKARVISNGDASNNYLRGYKDGSNRIYGLDGNDEIYGGA 810
QY 583 I---VNAGAGNDIIFVQGGKNIDGGDGHDRVYFSKDGFGNIT-----VDGTS 628
Db 811 LDDMLDGGAGNDMLSGGKADTLGGSGNDMLY-----GGTGNDTYRFAIGAGVDRIEESD 866
QY 629 ATEAGSYTVNRKVARGDYIYEVVKRQETKVGRTE--TIQY-----RDYEL 672
Db 867 AT-AGNTDVR-----PADVASSALTALERKDSDLVIKYGTSQDLTISNYFYSAEYKV 918
QY 673 RKVGYGYQST-DNLKSVBEVI--GSQFNDVFKGSKFND---IFHSGEGDDLLDGGAGDDR 726
Db 919 EQFTFSNGVTWDEAAIKARVISNGDANNYLVG--YNDGSRNIYGLDGNDTIYGGALDDM 976
QY 727 LFGGKGNRLSCDEGDDLLDGGSDVLLNGAGNDVYLFKRGDGNDTLYDG---TGN-DK 782
Db 977 LYGGAGNDMLSGGKADTLGGSGNDMLYGGTGNDTYRFAIGAGVDRIEESDATAGNTDV 1036
QY 783 LAFADANISDIEMIERTKEGIIVKRNDHSGSINIPIRWYITSLNQYQSNKTDHKEIQLIGK 842
Db 1037 VRFADVAFSALTALERKDSDLVIKYGTSQDLTISNYFYSA-----EYKVEQFTFS 1086
QY 843 DGSYITSQOI-----DKILQDKDGT 863
Db 1087 NG--VTWDEAAIKARVISNGDASNNYLRGYKDG 1118

RESULT 28
AB1905
outer membrane secretion protein alr0791 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AB1905
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, S.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB1905
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-993 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA072748.1; PID:gl7130136; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr0791

Query Match 6.6%; Score 305.5; DB 2; Length 993;
Best Local Similarity 22.7%; Pred. No. 5.6e-08;
Matches 205; Conservative 113; Mismatches 340; Indels 245; Gaps 41;

QY 121 SNVLSTLSSFLGTALAGIELDSLKKGAAPDALAKASIDLIN--ELIGNLSOSTQTIE 177
Db 27 TNVGTLSVDIIGNAISDPDASAL--KG-----IAVTFVDSNNGTWEYTLNNGTSWTFG 78
QY 178 AFSSQLAKLSTISQAKGFSNIGNKLO--NLNFSKTN--LGLLEIITGLLSISAGFA 230
Db 79 TPSLUTARLLP-----SNANTKIRFHPNANPSGTADINFYAWDQTTG-TSOSTANI- 128
QY 231 LADKNASTGKKAAGFELSNOVIGNVTKAISYVLAQVAAAG-----LSTTG 277
Db 129 LAGKGGTTA--FSTNYEGASITVTPVNDT-----APTLAGSPTLPTINEDAPLISNKG 180
QY 278 AVAALITSSIMLAISP---LAFNNAADKFNHA--NALDEFKQFRFEGYDGDHLLAEYQ 331
Db 181 SLLADLVRLGISDSDPDYQGIA-VTGADNNNGSQYSLDG-GVNNLNFNGNASDSATLLL 238
QY 332 RGVGTIEASLTITISTALGAVSAGVSAAGVSAVGTPIALL--VAGVTGLISGILEASKQA 389
Db 239 PSIRLYOGSLGSLPTSGWLKFGAS-----PPVPLFPVGGTOSLITGGIQLNSS 289
QY 390 MFESVANRLOGKILEMEKONGQNYFPDKGYDSRYAAYLAN-----NLKFLSEL 437
Db 290 IGSS-----GYSNY-----NSYAPILFNQAPPELDPVKGFTISFDVKI 327
QY 438 NKELEAERVIAITQORWNNIG--ELAGITKLERIKSGKAY-----ADAPE 482
Db 328 NGE-----THTSDDNGDGIQDRAGFSVIVTSDKTKAIELGFWTDEIWAQTASPLF 378
QY 483 DGKVEAGSNITLDAKTGIIDISNSGKKTQALHFTSPLL-----TAGTESR 529
Db 379 THSTTERAFRNTTAVTRYHLVVENNTYKLFAPDSSTPILSGNLRDYSAFNHSTAAPSPI 438
QY 530 ERLTNGKYSYINKLKFRGVKNWQVTDGEASSKLDFSKVIOQVVAETEGTDEIGLIIVNAKAG 589
Db 439 TSLFPDPYETPNFLFGD---NTTSAQASSNL-----TQVELQTNTRVRFVPENAD-- 485
QY 590 NDDIFVQGGKNIDGGDGHDRVYFSKDGFGNITVDGTSATEAGSYTVNRKVARGDYIYH 649
Db 486 ----YNGQANLTPRAWDGSGNAGVAGTTG--VNAAVNG-NATAFSSNTLTASITVSPINNP 538
QY 650 VVKRQETKVGRKTETIQYRDVELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIF 709
Db 539 I-----QGTGTL---DKLYGTANEDIINGNEGDYL 566
QY 710 HSGEGDDLLDGGAGDDRLFGKGNDRLSGDEGDDLLDGGSDGVLLNGAGNDVYIFRKG 769
Db 567 FGRAGNDLGGEGNDYLFGGTGNDTLDGGEGSDLLYGNEDNDIINGVGND--NLDCGT 624
QY 770 GNDTLYDGTGNDKLAFADANISDIEMIERTKEGIIVKRNDHSGSINIPIRWYITSLNQY-- 827
Db 625 GDDILRGTKGND--IYTVDTVGDVIEENPNNEG-TDKVNSYIS-----WTLGANLENLTL 675
QY 828 -----QSNKTDHKI-----EQLICKGSGSYIT-----SQID-KIILQD 858

Db 676 LGNTIIDGTGNELDHNNVAVNRLEGGDNDWLGKGNDDILIGNGNDRLNGETGED 735
 Qy 859 KKDGTV-----ITSQELKKLADENKSKQSLASDIASSLN-----KLVGSMALFGTANS 906
 Db 736 TLEGGGLGNDVVEIDSGVDVIEAADAAGIDTIVISSVDWTLGVNLENLTLVGNQATLGIGND 795
 Qy 907 VSS 909
 Db 796 LDN 798

RESULT 29
 G95851
 probable hemolysin-adenylate cyclase protein [imported] - Sinorhizobium meliloti (strain
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: G95851
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: G95851
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1072 <KUR>
 A:Cross-references: GB:ALU591985; PIDN:CAC48479.1; PID:g15139951; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid pSymB
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMD20079
 A:Genome: plasmid

Query Match 6.5%; Score 301.5; DB 2; Length 1072;
 Best Local Similarity 20.7%; Pred. No. 1e-07;
 Matches 207; Conservative 140; Mismatches 346; Indels 307; Gaps 43;

Qy 18 TKSLGNLYLAIPKDYDPQKGTGLNDFIKAADELGIAPLAEEPHNHTAKSVDTVNQFL 77
 Db 152 TISNFENVY-----GGSGNDILTGDDRSNILR-GEAGNDILGGGADDDLLS--- 196

Qy 78 SLTQTGIAISATKLEFLQKSTNKLAKGL-----DSVENIDRKLKASN 122
 Db 197 GGAGNDTADGGTGIDTDFLREKTSVVVQLSGANAATVFGGVAEDTIRNVNIVGGTAD 256

Qy 123 VLSTLS-SFLGTALAGIELDSLKKGDAA-----PDALAKASIDL-----I 162
 Db 257 --DTLSGDAAANKLSARGNDWLKGGGADTLGGEDSDTADYDKAAIAVALANGNPV 314

Qy 163 NEIIGNLSQSTQTIEAFSSQLAKGLSTISQAKGFSNIGNKLNQNFKNLGLIITGLL 222
 Db 315 TVTVGGIAEDL-----IAKNIENIVGGSGDITIGDAAANA--PRGGLGADVLDGGG 363

Qy 223 SGISAGFALADK-----NASTGKKVAAGFELSNOV--IGNVTKAISSVVLQORVAAG 272
 Db 364 GSDTADF--SDKVQSVVLALNGAVDAIAAAGGTAEADTVRNIENTITGGSGNDQFTGDAAN 421

Qy 273 LSTTGAVAALI-----TSSIMLAISPLAFMNAADKFENHANALDEFAKQFRK 318
 Db 422 TFRGLGADVLDGGGSDTADYDKTTSVVVT-----LAGANPTTAFVDMGADSLSRNENI 478

Qy 319 FGYGDDHLLA--EYQR-----GVGTIEASLTITISTALGAVSAGVSAA--AVGSVGTPTIAL 370
 Db 479 IGGSGNDVLGDDGFQNVLDGGAGTDMADYSASAKGIAVMLNGANDAKVIVGSAEDTLR- 537

Qy 371 LVAGVTG-----LISGILEASKQAMFESVANRLOG-----KILEWEKONGKONGYFDKGYDS 421
 Db 538 NIENVTGSAPADVTG-----DAQNILLGGSGDIL---KDDGGQDVIDGAGT 584

Qy 422 RYAAYLANNLKLFLSELNKELEAERVIAITQORWNNNIGELAGITK-----LGE 469
 Db 585 DTADF-----SEKTAAVVLALAGAA--NAIATVGLAEDTVRNIESIFGGIGA 630

Qy 470 RIKSKAYADAPEDGKKVBEAGSNITIDAKTGIIDISNSNGKKTQALHFTSPLTAGTESR 529
 Db 631 DVLTDGNSNTIRGG-----AGAD-SLDGGAGVDTVDRD--KTKSVAITLDGATFVT--- 680

Qy 530 ERLTNGKYSYINKLKFG-----RVKNWQ-VTDGEASSKLDKDFKVIQORVAETEGDEIGLI 583
 Db 681 -----VKVGGVIEDTIRNFENISGSGAGDMLT-----GDGLANV 714

Qy 584 VNAKAGNDIIFVGGKKNIDGGGHRVYFSKDGFGFNITVDGTSA----- 629
 Db 715 LVGNDGADTLRGGLGKVDLDGGNGVDADYLEKTDAISVTLNGTASAAVLVGGTAEDTIR 774

Qy 630 -----TEAGSYTVNRKVARGDYIHEVVKQETKVKKRTETIOYRDYELKVG----- 676
 Db 775 GVENILSGSGADTLVGDTA--SNMFRGALGADFIDGGAGVDTDADYRE----KTGSVEVALS 829

Qy 677 -----YGYQSTDNLKSVEEVIQSQFND-----VFKGSKFNDIFHSGEGDDLLD 719
 Db 830 GASDSFVFGVVEDTIRNIENVFGKGNLTLTGDLGLANTLNGDGLLTTGGGADILD 889

Qy 720 GGAGDD-----RLFGKGNDRLSGDEGD 742
 Db 890 GGAASDTASVYRDKTASVSVTLDCATYTTVTGVAEAEDTIRNVENIMGTGNDLSLGDANA 949

Qy 743 DLLDGGSGDDVLNGGAGNDVYIFRKGDN---DTLYDGTGNDKL----- 783
 Db 950 NLLSGGSGSILFGGAGADIFQDFALGSTNVDTVLDFDTAGDRFLSKSIFTSLSGGTLA 1009

Qy 784 -----AFADA-----NISDIMERTEKGIIVKRNDSHGSIN 814
 Db 1010 ATQFYAAADATAAQNQVNIYYTTSGLY--YDADGSL 1047

RESULT 30

AG2137

hypothetical protein all12654 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AG2137

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1417 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA074353.1; PID:g17131747; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all12654

Query Match 6.2%; Score 288; DB 2; Length 1417;

Best Local Similarity 21.6%; Pred. No. 7.5e-07;

Matches 219; Conservative 131; Mismatches 355; Indels 310; Gaps 50;

QY 790 ---ISDIMERTKGIIVKRNDS-----GSIIPRWIITSNLQNY 827
Db 695 VDSIGDYVLENAOGTDLVQSSISYITLNSLENLTLCGTSAINGTGNKLNVIITGNSGN 754
QY 828 QSNKTDHK-----IFQLGKDGSYI-----TSQIDKIL 856
Db 755 TLNGGDGNDTLNGSAGVDTLLGGNGDILVGGTGNDTLTCGVGRDRFTFNSRSEGDIT 814
QY 857 Q-DKKGDTVITSQELKKLADENKSKQSASDIASL---NKLVGSMALFGTANVSNN 910
Db 815 DFNVDVITV-----VSAAGFGGLVVGAAIASQFLGSAATTASH 856

RESULT 34
G87572
calcium-binding protein [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: G87572
R.Nieman, W.C.; Feildblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: G87572
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-650 <STO>
A:Cross-references: GB:AE005673; NID:gl3424183; PIDN:AAK24579.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2610

Query Match 5.9%; Score 276; DB 2; Length 650;
Best Local Similarity 26.0%; Pred. No. 1e-06;
Matches 107; Conservative 51; Mismatches 118; Indels 136; Gaps 17;
QY 588 AGNDIFVGGKMNIDGGHGRVYSKDGFGNITVDGTSATEAGSYTVARKVARGDIY 647
Db 34 AGDDAIAAGHEGNDVLQGGDGDYML-----GGFGNDTLDGAGSDWAAY----- 77
QY 648 HEVVKRQETKVKRTETIQRDEL---RKVGYGYSQSTNLKSVVEVIGSFNDFVFKGSK 704
Db 78 -----EDATAGVKV-----DLNLSGAQNTGGG---GTRLTSLNLYGSAFNDTLIGNA 123
QY 705 FNDIFHSGEGDLDLGGAGDRLFGCKGNDRLSGDEGDDLLDGGSGDDVLNGAGND--- 761
Db 124 GDNMITGAGADSITGKGDDTLGWSAGNDTLDGGDGMVVGAGDDVIKGGAGDWS 183
QY 762 -----VYIFRKG-DGN 771
Db 184 YEDATAGVTVDLTKTTAQDTIGAGKDTLGSVENLWGSKFADVLTGDAGSNYIW--GDAGD 241
QY 772 DTLYDGTGNDKLA--FADANISD-----IMERTKEGIIVKRNHSSGSIINPRWITSN 823
Db 242 DKLYGAGDGFAGAGVNVIDGGEGFTDIYGMGAVG-----EVDLSRIATSR 292
QY 824 LQNYQSNKTDHRIEQLIK-----DGSYITSQIDKILQDKDGTWITSQELKL 873
Db 293 FGDTSISDTLSSIELVTGSIYADFIIIGNMAENLYGDAGNDVLRAGEGDVLGEGE--- 348
QY 874 ADENKSKLASDIASSNLKLVG-----SMALFGTANSVSS-----NALQPI 916
Db 349 GDDFLIGSL--SDI-----LTGGGDDQLSVSGGSNYVDGGDGVDAQTFT 392

RESULT 35
H95964
probable outer membrane secretion protein SMB21543 [imported] - Sinorhizobium meliloti
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: H95964

R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9899-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: H95964
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1112 <KUR>
A:Cross-references: GB:AL591985; PIDN:CAC49384.1; PID:gl5140870; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMB21543
A:Genome: plasmid

Query Match 5.8%; Score 269.5; DB 2; Length 1112;
Best Local Similarity 24.5%; Pred. No. 4.7e-06;
Matches 127; Conservative 56; Mismatches 143; Indels 193; Gaps 24;
QY 407 KONGQNYF---DKGYDSRYAAYLANLKLSELNKELEA---ERVIAL-----TQORW 454
Db 183 QQYSVMYFDGIDVGETSRWSA-----STPLMADIEAVIRRFSTVDENGVRTTYQHI 234
QY 455 DNNIGELA---GITKLGERIKSGKAYADAFEDGKVEAGSNITLDAKTGIIDISNSGKK 511
Db 235 DLNTGDNVYFGSTQYGYQITS-----SGHQHIGFAIHDTGGVTDIDFSGS--- 281
QY 512 TQALHFTSPLLTAGTESRRLTNGKYSYINKLFGKRVKNQVTDGEASSKLDFSKVIQRV 571
Db 282 -----TAGTIL--DLRAGQFSSVN---GHSNNVSI PAGHNADAADY----- 317
QY 572 AETGTDDEIGLIVNAKGN--DDIFVG--QGKMNIDGGDHRVYSKDGFGNITVDGTS 628
Db 318 -----YIETGIGSRFDDILIGNDANTLDGRSGGDRM--AGNGGDDTYFVDSL- 363
QY 629 ATEAGSYTVARKVARGDIYHEVVKRQETKVKRTETIQRDELKRVKG-----YGVQST 682
Db 364 -----EDIVREANGNDTVILLRNKLRKIRKIANVENIYADEST 403
QY 683 -----DN--LKSVEEVIGSFNDFVFKGSKFNDFIHSGEGLDLDGGA----- 722
Db 404 TQPGGGGASPPSAGDNTWTSI--IFGCGNDTLDGGAGDDTIFGKGGDILLIGRDSLAS 461
QY 723 -----GDDRLFGKGNDRLSGDEGDDLLDGGSGDDVLNGAGNDV 762
Db 462 RDINNTIDVEDLEQTESDDGNDTLYGGGNDTTLGGQGNIDLDGGAGDVLGQDGV- 520
QY 763 YIFRKGNDTLYDGTGNDKLAFLADANISDIIMERTKEG-----IIVKRNHSSGSIINPR 817
Db 521 -IFRGGAGVDT-----VDFSKESPQLLVNLTATNVASSGT----- 554
QY 818 WYITSLNQYQSNKTDHRIEQLIKGKGSYITSDQIDKIL 856
Db 555 -----ASGDTFHSIENLIGSD-----DRIDRFI 577

RESULT 36
G95405
hypothetical protein Sma2111 [imported] - Sinorhizobium meliloti (strain 1021) magaplas
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: G95405
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, K.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A;Title: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti*
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: G95405
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-539 <KUR>
A;Cross-references: GB:AE006469; PIDN:AAK65809.1; PID:g14524311; GSPDB:GN00165
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; Hyman, R.W.; Jones, T.L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, H.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: Sma2111
A;Genome: plasmid

Query Match 5.8%; Score 268; DB 2; Length 539;
Best Local Similarity 26.4%; Pred. No. 2e-06;
Matches 103; Conservative 59; Mismatches 160; Indels 68; Gaps 13;

Qy 519 SPLLT--AGTRESRLTNGKYSYINKLKFGKRVKMQVTDGEASSKLPFSKVIQVRAETEG 576
Db 21 SPIMTRIVGTSADTLOG-----MAEGDRIWSLDGNDVVDGGA---G 59

Qy 577 TDEIGLVNAKAGNDDIFVGGKKNIDGGDGHVFFY-----SKDGGFGNITVDGTSATE 631
Db 60 DD-----FVDCGAGDALTSSGDFEFTGEGNDRLSFIVGGAARGGTGVTLVGYAAI 115

Qy 632 AGSVT-----VNRKVARGD-----IYHEVVRQRTKVKRTTETIQYRDYELRKV--GYGY 679
Db 116 SDAFLFCGMHGAFAAGDLGVKGNHLYFLDIERLSLTTCIGDDRIIATGFSFVHVHTGAGD 175

Qy 680 QSTONLKSVERVIGSQNFQNVFKSKFNDIFHSGEGDLDLGGAGDRLFGKGNDRLSGD 739
Db 176 DRVETGTDGDDQIYAGDGRDLFGGGGDDFTHTGGDDYVNGNDRLREGEDGNDLSVGG 235

Qy 740 EGDLLDGGSGDDVNLGAGNDVYIFRKGNDTLTYDGTGNDKL--AFADANI-----S 791
Db 236 RGNDRLDGGSGDDVNGCGDND--SLTGLGSDVTVTGAGDDYLSNFSFAGDILLGGDGN 293

Qy 792 DIMIERTKEGIIVKRNHSGSINIPRWYITSNLQVQSNKTDHKIEQLKDGSGYITSDQ 851
Db 294 DTLSAGGEDTAYGWSDLVGGAGDRLHI-----YTDGSLGALDGGEG----FDR 339

Qy 852 IDKLQDKCGTGVITSQELKKLADENKSOQ 881
Db 340 ASIALDDVSAGFVLDASRFGSIEEFNITVK 369

RESULT 37
E95995
hypothetical calcium binding protein [imported] - *Sinorhizobium meliloti* (strain 1021)
C;Species: *Sinorhizobium meliloti*
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: E95995
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernar
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: E95995
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-387 <KUR>
A;Cross-references: GB:AL591985; PIDN:CA49629.1; PID:g15141116; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; Hyman, R.W.; Jones, T.L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, H.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: Smb21402
A;Genome: plasmid

Query Match 5.8%; Score 267.5; DB 2; Length 387;
Best Local Similarity 30.8%; Pred. No. 1.3e-06;
Matches 89; Conservative 39; Mismatches 92; Indels 69; Gaps 15;

Qy 545 FGRVKWQVTDGEASSKL-----DFSKVIQVRAETEGTDEI--GL-----IV 584
Db 31 YGGVGNDRIFGGDGNDIYVGGVDYAE-----GEGNDRLDGGLGNDAPFMGIGNDIF 84

Qy 585 NAKAGNDDIFVGGKKNIDGGDGHDRVFFYFKGCGFNITVDGTSATE-----AGSVTVNR 639
Db 85 DGGEGNDVLDGGHGHQMLGGRGDKIF---CGAGEDYIDADEGDDIIFAGSGDDGFNN 140

Qy 640 KV--ARGDIYHEVVRQRTKVKRTTETIQYRDYELRKVGYGYOSTNLKSV---EEVIGS 694
Db 141 RVNPATCKLTQQAAGV---GAGNDTI-----YEGGNDALKGQSGNDRVYGG 184

Qy 695 QFNDVFKSGKFNDIFHSGEGDLDL-----GGAGDRLFGKGNDRLSGDEGDDL 745
Db 185 IGGDIYVGGNGNLYDGGDGNVDLSEGGSEAHGSGGNDRIISVAGNDRAFDGGDDIL 244

Qy 746 DGGSGDDVNLGAGNDVYIFRKGNDTLTYDGTGNDKLAFADANI 794
Db 245 TSAGAGDDYLSGGIHQD-RLF-GGTGNDTLRGDAGRVLS-GDAG-SDIL 289

RESULT 38
AD2263
hypothetical protein alr3659 [imported] - *Nostoc* sp. (strain PCC 7120)
C;Species: *Nostoc* sp. PCC 7120
A;Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AD2263
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD2263
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-589 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA75358.1; PID:g17132792; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr3659

Query Match 5.7%; Score 266.5; DB 2; Length 589;
Best Local Similarity 25.7%; Pred. No. 2.7e-06;
Matches 114; Conservative 64; Mismatches 147; Indels 119; Gaps 21;

Qy 411 GQNVFKGYDS-RYAAYLANNLKFLSELNKELEAERVIAITQORWNNIGELAGITKLGE 469
Db 36 GDDYLYGLDNDKLYGGLGNDKLYGGEGDNLNGE-----QGNNSLYGEAGDDTLDA 87

Qy 470 RIKSGKAYADAFEDGKKV---EAGSNITLDAKTG--IIDSNSNGKKTQALHTFSPLLTA 524
Db 88 GNSTGNLLDGGEGNDEIYVYSSGNNTLKGVGNDILDVRFSGKN-----ILDG 138

Qy 525 GTSRE---RLTNKGYSYNKLKFGRVKMQVTDGEGASSKL-----DFSKVIQVRAETEG 577
Db 139 GSGNDQLYARYSQGN-----NLRGGSGDDYLVGLDLSAGNNVLNGDGDNDILDSYQWKGK 194

Qy 578 DEIGLVNAKAGNDDIFVG---QCKWIDGGDGHDRVFSK-----DGGF 619

Db 195 N-----IVSGNGN-DIFVAYGQANTLNGSGDDSFYISPSADTVVHPLVTVQVGGI 249
Qy 620 GN--ITVDGTSATEAGSYVNRKVARGDIYHEVVKRQETKVKRTEFIQYRDYELRVGY 677
Db 250 GNDVLYIDYSTAIAGITTSFNTTKQGLI-----TADTNQVR-----286
Qy 678 GYQSTDNLKSVVEVIGSFQNFVFKSGKFNDFHSGE-GDDLDDGAGDDRL-----727
Db 287 -YKNIERL-----NIIGTSHDDIIVGNSDDIIDGSGGNDTLGAGGNDIILIRDSRSH 341
Qy 728 ---FGKGNDRLL--SGDEGDDLLDGGSGDDVLN-----CGAGNDVYIFR 766
Db 342 NTVYGGAGNDSLYANGTAGTNLGGDNDYLSAKNYSYFNASQTLIGAGGDDTLDSV 401
Qy 767 KGDGNDTLYDGTGNDKLAADANI 790
Db 402 GSYGNFLYGGAGNDFL-YADYSV 424

RESULT 39
AD1841
hypoetical protein alr0276 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AD1841
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD1841
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2348 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAB77800.1; PID:gl7135254; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr0276

Query Match 5.7%; Score 266; DB 2; Length 2348;
Best Local Similarity 20.4%; Pred. No. 2.1e-05;
Matches 193; Conservative 121; Mismatches 332; Indels 298; Gaps 38;
Qy 16 NSTKSLKML-----YLAIPKDYDPKQGTGLNDFIKAADELGIARLAEENHETAKK--S 69
Db 1479 NSAFGGIRTIEGVVVAAPP-----GSGGLNGFFVQEE-----ADTDNDSTTSEGIFV 1526
Qy 70 VDTVNOFLSLTQTGIAISATKLEKFLQKHSNKLAKGLDSVENIDRKLKGNVLSLTS 129
Db 1527 FDTPTGQF-----SGSVGDKVRVTGVSFSTNGVSSLTQLSSVSSINLADILPRMSN 1581
Qy 130 FL--CTALAGIELDSLIIKKGAAPDALAKASIDILNIEIGLSOSTQTIEAFSSQAKLG 187
Db 1582 IQLPTVTADLRYEGMR-----VNISAGN-GDLTIVT-EHF--QLGRFG 1621
Qy 188 STISOAKGFSNIGNKLQNLNFSKTNLGLLEITGLLSISAGFA-----LADKNA 236
Db 1622 QVLSATGTSNPG-----TDGRLEQYTFQNDPSVAGVAAVLDIAKRRRIILDDGS 1672
Qy 237 ST-----GKVAAGFELSNOVIGNTKAISSVVLAQR-----VAAGLSITGA 278
Db 1673 STQNPAITIFGRGGPLSA-----TNTLRGGDTVASITGVLDQRFEGYRVQVSTGVDFTPA 1728
Qy 279 VAALITSSIMLAISLAFNAAADKEN-----HANALDEFKQFRF-----319
Db 1729 NRPATTPDVGTLKVASFNVLNLYENGDTGSGFTSPQRGAENLTFNRQREKTIAAIL 1788
Qy 320 -----GYDGDHLLABYQRGVGTIEASLTITSTALGAVSAGVSAAGVAV 364
Db 1789 GLNADWVGLIETENDGYGANSIQDLINGLNAVAGAGTYAFINPGLSQLGTDIAVGF-FI 1847

Qy 365 GTPIALIVAGVTGLISGILEASQKAMPESVANRLQKILEWEKQNGQNFYDKGYDSRYA 424
Db 1848 YKNSVTPVGVAAATVA-----DGFQOGAFDN-----1873
Qy 425 AYLANNLKFLSELNKELEAERVIAITQORWDNNIGELAGITKIGERIKSKAYADAFEDG 484
Db 1874 ----NNRKPLAQTPRQ-----NSTGE--QFTAVINHFKS-----1901
Qy 485 KVEAGSNITLDAKTIIDISNSNGKKTQALH-----FTSPILLTACTE-----527
Db 1902 KGSSSGNPGDADAGD---QGLSNGTRTRASQDLAAWLATNPCTGTTDTDTYLLGLDLNAYA 1958
Qy 528 -----SRERLTNGKYSINKLKFRVKVN-----WQVTDGEAS 559
Db 1959 QEDPIRALENAGYNNLLPNTYTYVFGQGALDHALANASLASQSVSTAVKWHI--NADEP 2017
Qy 560 SKLDFSKVIQVRAET-----EGTDE-----IGLIVNAK--AGND-----DIF 594
Db 2018 NVLDYNTNFKSVGQOTSLSYSPDAFRSSDHDPFVIGLNLNTAPIAVNDIATNTNENTAVNIN 2077
Qy 595 VGQGNMIDGGDGHDRVYFSKDGFGNITVDGTSATEAG---SYTVNRKVARGDIYHEV 651
Db 2078 VLTNDSVNGDSLQLSLVSNPVGAVVNDNGTGFNFADDFITTPNTGYLNGDSFTYSI 2137
Qy 652 KROETKVKKRTETIQYRDYELRVKVGYGQSTDNLKSVEEVIGSFQNFVFKSGKFNDFHS 711
Db 2138 --SDKKGATATVSL-----TINASGGIIGTPDNDILATGTNRNDLIRG 2179
Qy 712 GEGDLDLGGAGDRLRFGKGNDRLSGDEGDDLLDGGSGDDVNLGAGNDVYIFPRKGDGN 771
Db 2180 LGGNDLLIGCGNDTLYGDRGDKILLGCGNDTLYGCGDDNDTLTGCGDDLLV--GGKGN 2237
Qy 772 DTLYDGTGNDKLAADAN-----ISD-----IMERTKEGI 802
Db 2238 DLLTGGNGRDRFYLSDTRTGTGEFDITDFKVGQDTILIPRAEFGL 2281

RESULT 40

E95933
probable calcium-binding exported protein [imported] - Sinorhizobium meliloti (strain 10.1)
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: E95933
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan, P.; Finan, T.M.; Weidner, S.; U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo-
symbiont Sinorhizobium meliloti
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95933
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-387 <KUR>
A:Cross-references: GB:AL591985; PIDN:CAC49133.1; PID:gl5140618; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, L.; Hyman, R.W.; Jones, T. Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo-
symbiont Sinorhizobium meliloti
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Smb21229
A:Genome: plasmid

Query Match 5.7%; Score 263.5; DB 2; Length 387;
Best Local Similarity 25.4%; Pred. No. 2.1e-06;
Matches 99; Conservative 51; Mismatches 111; Indels 129; Gaps 15;
Qy 546 GRVKNQVQVTEGEASKLDFSKVIQVRAETGTEIGLIVNAKAGNDIIFVCGKKNIDGG 605
Db 77 GGLGNDILDDGSGSDILDGG-----DGHDOI-----LGGRGHDKIFGAGGEYIDAG 123

| | | | | | | | |
|----|-----|---|-------------------------|----------------|-----------------|----------|-----|
| Qy | 606 | DGHRVFY-SKGGFGNITVDGTSATEAGSYTVNRKRVARGDIYHVVKRQETFKVGR | TET | 664 | | | |
| Db | 124 | DGDDIIVAGSGDDGFNNRVPATG | -----QVTQAVAGGAGHDKI | 163 | | | |
| Qy | 665 | IQYRDYELRKVGYGYQSTDNLKS | V---REVIGSQFNEDVPFGSKFN | DI-FHSGEGD | LLD--- | 719 | |
| Db | 164 | -----YBEGNDALKGSGDNWVYGGI | GNIDIVDGGDGNLYLGGDGN | DVL | DSE | 211 | |
| Qy | 720 | -----GGAGDRLFGKGNDRLSGD | DDLDCGSGDDVLNGCAGNDVY | IFRKG | DND | 772 | |
| Db | 212 | GGIDEAHGGTKIIVGAGD | RVFGDDGDDMLYGEAGN | DSLGG | NVYD-RIF-GGD | 269 | |
| Qy | 773 | T-----LYDGTGNDKLAFADA | -----NISD | MIERTKE | 800 | | |
| Db | 270 | TLRGDAGRDVLIGEAGTDILWGGAD | SDRFVFKGAPVLSQDVT | MD | FQDGVDFLVIEK--- | 326 | |
| Qy | 801 | GIIVKRNDHSGSINIPRWYIT | SNLQNYQSNKTDHKIEQLIGK | GSYITSDQID | IKDLQDK | 860 | |
| Db | 327 | -LGIKQYSSGA-----AGTV | VAYDATG | GDVLVK-----GYD | SSGKA | FILLSDDP | 370 |
| Qy | 861 | DGTWITSQELKLADENK | SQKLSAD | DIASS | 890 | | |
| Db | 371 | NGT-----LSAN | FSSS | 382 | | | |

Search completed: February 17, 2004, 10:12:18
Job time : 40 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:09:49 ; Search time 45 Seconds
(without alignments)

5315.884 Million cell updates/sec

Title: US-10-069-799-5

Perfect score: 4647

Sequence: 1 MSINIVKSNIQAGLNSTKS.....SSNALQPIQTQTGILAPSV 927

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------|--------------------|
| 1 | 4642 | 99.9 | 927 | 2 Q93G12 | Q93G12 moraxella b |
| 2 | 2358.5 | 50.8 | 946 | 2 Q9EV24 | Q9EV24 mannheimia |
| 3 | 2356 | 50.7 | 953 | 2 Q9EV23 | Q9EV23 mannheimia |
| 4 | 2355 | 50.7 | 953 | 2 Q9ETX2 | Q9ETX2 mannheimia |
| 5 | 2355 | 50.7 | 953 | 2 Q9EV25 | Q9EV25 mannheimia |
| 6 | 2345 | 50.5 | 953 | 2 Q9EV26 | Q9EV26 mannheimia |
| 7 | 2338.5 | 50.3 | 953 | 2 Q9EV32 | Q9EV32 pasteurella |
| 8 | 2336.5 | 50.3 | 953 | 2 Q9EV33 | Q9EV33 pasteurella |
| 9 | 2332.5 | 50.2 | 953 | 2 Q9EV28 | Q9EV28 pasteurella |
| 10 | 2332 | 50.2 | 953 | 2 Q9ETG5 | Q9ETG5 pasteurella |
| 11 | 2329.5 | 50.1 | 953 | 2 Q9EV31 | Q9EV31 pasteurella |
| 12 | 2329.5 | 50.1 | 953 | 2 Q9EV27 | Q9EV27 pasteurella |
| 13 | 2329 | 50.1 | 953 | 2 Q9EV29 | Q9EV29 pasteurella |
| 14 | 2326 | 50.1 | 953 | 2 Q9EV34 | Q9EV34 pasteurella |
| 15 | 2320 | 49.9 | 954 | 2 Q9EU01 | Q9EU01 pasteurella |
| 16 | 2307 | 49.6 | 953 | 2 Q9EV30 | Q9EV30 pasteurella |

Q9ev22 pasteurella
Q9eud4 pasteurella
Q93np1 actinobacil
Q8kwz6 actinobacil
Q8kwz9 actinobacil
Q9rcg8 pasteurella
Q93np0 actinobacil
Q8fe01 escherichia
Q8ga40 escherichia
Q8g924 escherichia
Q43892 actinobacil
Q85101 escherichia
Q9lc58 escherichia
P71223 escherichia
Q46716 escherichia
Q93np2 actinobacil
Q47461 escherichia
Q47262 escherichia
Q91469 bordetella
Q51865 pasteurella
Q9pel7 xylella fas
Q9pf19 xylella fas
Q9pat8 xylella fas
Q51868 pasteurella
Q937v6 bordetella
Q9jtu8 neisseria m
Q937w0 bordetella
Q8y2t5 ralatonia s
Q932y7 bordetella

ALIGNMENTS

RESULT 1

Q93G12 ID Q93G12 PRELIMINARY; PRT; 927 AA.
AC Q93G12; STRAIN=Tifton I;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RTX toxin.
GN MBXA.
OS Moraxella bovis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
OX NCBI_TaxID=476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tifton I;
RX MEDLINE=21388402; PubMed=11497442;
RA Angelos J.A., Hess J.F., George L.W.;
RT "Cloning and characterization of a Moraxella bovis cytotoxin gene."
RL Am. J. Vet. Res. 62:1222-1228(2001).
DR EMBL; AF205359; AAK84651.1; .
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysin_cbind; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00018; EF_HAND_1.
DR PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ SEQUENCE 927 AA; 98845 MW; F4B703577E10A96D CRC64;

Query Match 99.9%; Score 4642; DB 2; Length 927;

Best Local Similarity 99.9%; Pred. No. 1.6e-199;

Matches 926; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSINIVKSNIQAGLNSTKSGLNLYLAIPKDYDPQKGGTINDFIKAADELGIARLAEEP 60

Db 1 MSINIVKSNIQAGLNSTKSGLNLYLAIPKDYDPQKGGTINDFIKAADELGIARLAEEP 60

```

Qy 61 NHTETAKSVDTVNOFLSLTGTGTATISATKLEKFLQKSTNKLAKGLSDSVENIDRKLGKA 120
Db 61 NHTETAKSVDTVNOFLSLTGTGTATISATKLEKFLQKSTNKLAKGLSDSVENIDRKLGKA 120

Qy 121 SNVLSTSSFLGTALAGLELSDLIKKGDAAPDALAKASIDILNEIIGNLSOSTOTIEAFS 180
Db 121 SNVLSTSSFLGTALAGLELSDLIKKGDAAPDALAKASIDILNEIIGNLSOSTOTIEAFS 180

Qy 181 SOLAKLGSTISQAKGFSNIGNKLQNLNFKSNLNLGLEIITGLSGISAGFALADKNASTGK 240
Db 181 SOLAKLGSTISQAKGFSNIGNKLQNLNFKSNLNLGLEIITGLSGISAGFALADKNASTGK 240

Qy 241 KVAAGFELSNQVIGNVTKAISYVLAQRAVAGLSTTGAVAAALITSSIMLAISPLAFMNA 300
Db 241 KVAAGFELSNQVIGNVTKAISYVLAQRAVAGLSTTGAVAAALITSSIMLAISPLAFMNA 300

Qy 301 DKFNHANALDEFKQFRKFGYDGHLLAEYORGVTIEASLTITSTALGAVSAGVSAAV 360
Db 301 DKFNHANALDEFKQFRKFGYDGHLLAEYORGVTIEASLTITSTALGAVSAGVSAAV 360

Qy 361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQKILEWEKQNGQNYFDKGYD 420
Db 361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQKILEWEKQNGQNYFDKGYD 420

Qy 421 SRYAAYLANNLKFLSELNKELEAERVIAITTOORWNNIGELAGITKLGERIKSKAYADA 480
Db 421 SRYAAYLANNLKFLSELNKELEAERVIAITTOORWNNIGELAGITKLGERIKSKAYADA 480

Qy 481 PEDGKKEVAGSNITLDATGTTIDISNSNGKKTQALHFTSPLLTAGTESRRLTNGKYSYI 540
Db 481 PEDGKKEVAGSNITLDATGTTIDISNSNGKKTQALHFTSPLLTAGTESRRLTNGKYSYI 540

Qy 541 NKLKFRVKNQVTDGEASSKLDPSKVIQVAETEGTDEIGLIVNAKAGNDDIFVGGQKM 600
Db 541 NKLKFRVKNQVTDGEASSKLDPSKVIQVAETEGTDEIGLIVNAKAGNDDIFVGGQKM 600

Qy 601 NIDGGDGHDRVFSKQDGFNITVDGTSATAGSVTVNRKVARGDIIYHEVVKROETKVGK 660
Db 601 NIDGGDGHDRVFSKQDGFNITVDGTSATAGSVTVNRKVARGDIIYHEVVKROETKVGK 660

Qy 661 RTETIQYRDYELRKVGYSQSDNLKSVVEVIGSQFNDVFKGSKFNDFHSGEGDLDLDG 720
Db 661 RTETIQYRDYELRKVGYSQSDNLKSVVEVIGSQFNDVFKGSKFNDFHSGEGDLDLDG 720

Qy 721 GAGDRLFGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDNNDTLYDGTGN 780
Db 721 GAGDRLFGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDNNDTLYDGTGN 780

Qy 781 DKLAFADANISDMITERTKEGIIKVRNDHSGSINIPRWYITSNLQYNSKTDHKIBOLI 840
Db 781 DKLAFADANISDMITERTKEGIIKVRNDHSGSINIPRWYITSNLQYNSKTDHKIBOLI 840

Qy 841 KGQGSYITSDQIDKILQDKDGTVITSELKKLADENKSKLSASDIASSLNKLVGSMAL 900
Db 841 KGQGSYITSDQIDKILQDKDGTVITSELKKLADENKSKLSASDIASSLNKLVGSMAL 900

Qy 901 FGTANSVSSNALQPIQTQTGILAPSV 927
Db 901 FGTANSVSSNALQPIQTQTGILAPSV 927

RESULT 2
Q9EV24 PRELIMINARY; PRT; 946 AA.
AC Q9EV24;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Mannheimia glucosida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

```

```

OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=95401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH574;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT haemolytica.";
RL J. Bacteriol. 183:1394-1404 (2001).
DR EMBL; AF314521; AAG40305.1; -.
DR InterPro; IPR001753; EnCoA_hydrtse.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNDRPT.
DR PROSITE; PS001488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 946 AA; 101480 MW; 25C077858BDC76C4 CRC64;

Query Match 50.8%; Score 2358.5; DB 2; Length 946;
Best Local Similarity 49.9%; Pred. No. 1.9e-97;
Matches 471; Conservative 182; Mismatches 258; Indels 33; Gaps 15;

Qy 1 MSNINV-IKSN---IQAGINST---KSGLNLYLAIPKD--YDPQKGGTINDIFIKADE 50
Db 1 MSNINV-IKSN---IQAGINST---KSGLNLYLAIPKD--YDPQKGGTINDIFIKADE 50
Qy 5 LTNITNWLKSSWLTAKSGLNRTGOSLKTGAKKILYIPKQYQYDTEKNGQLDLVKAABE 64
Db 5 LTNITNWLKSSWLTAKSGLNRTGOSLKTGAKKILYIPKQYQYDTEKNGQLDLVKAABE 64
Qy 51 LGTARLAEEPNTHTAKKSVDTVNOFLSLTGTGTATISATKLEKFLQKSTNKLAKGLSDSV 110
Db 51 LGTARLAEEPNTHTAKKSVDTVNOFLSLTGTGTATISATKLEKFLQKSTNKLAKGLSDSV 110
Qy 65 LGTEVQKEBNDTAKAQTSLGTQNVGLTGERGIVLSAPQLDKLQK---TKVQQAATGSA 121
Db 65 LGTEVQKEBNDTAKAQTSLGTQNVGLTGERGIVLSAPQLDKLQK---TKVQQAATGSA 121
Qy 111 ENIDRKLGKASNVLTSLSTSLFELGTALAGLELSDLIKKGDAAPDALAKASIDILNEIIGNLS 170
Db 111 ENIDRKLGKASNVLTSLSTSLFELGTALAGLELSDLIKKGDAAPDALAKASIDILNEIIGNLS 170
Qy 122 ENLTGKFSNAKTVLSGIQSLGSLVAGWDLDEALQK-NSNELTLAKAGLELTNSLIENIA 180
Db 122 ENLTGKFSNAKTVLSGIQSLGSLVAGWDLDEALQK-NSNELTLAKAGLELTNSLIENIA 180
Qy 171 QSTQTIETAFSSQAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLIITGLSGISAGF 229
Db 171 QSTQTIETAFSSQAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLIITGLSGISAGF 229
Qy 181 NSVTDLDAFDQINQLGSKLQNVKLSGLGKLSGDKTSLGSLDWSGLSGATAAL 240
Db 181 NSVTDLDAFDQINQLGSKLQNVKLSGLGKLSGDKTSLGSLDWSGLSGATAAL 240
Qy 230 ALADKNAKSTGKVAAGFELSNQVIGNVTKAISYVLAQRAVAGLSTTGAVAAALITSSIML 289
Db 230 ALADKNAKSTGKVAAGFELSNQVIGNVTKAISYVLAQRAVAGLSTTGAVAAALITSSIML 289
Qy 241 VLADKNAKSTGKVAAGFELSNQVIGNVTKAISYVLAQRAVAGLSTTGAVAAALITSSIML 300
Db 241 VLADKNAKSTGKVAAGFELSNQVIGNVTKAISYVLAQRAVAGLSTTGAVAAALITSSIML 300
Qy 290 AISPLAFMNAADKFNHANALDEFKQFRKFGYDGHLLAEYORGVTIEASLTITSTALG 349
Db 290 AISPLAFMNAADKFNHANALDEFKQFRKFGYDGHLLAEYORGVTIEASLTITSTALG 349
Qy 301 AISPLAFAGIADKFNHANALDEFKQFRKFGYDGHLLAEYORGVTIEASLTITSTALG 360
Db 301 AISPLAFAGIADKFNHANALDEFKQFRKFGYDGHLLAEYORGVTIEASLTITSTALG 360
Qy 350 AVSAGVSAAGVSAAGVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQKILEWEKQ 409
Db 350 AVSAGVSAAGVSAAGVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQKILEWEKQ 409
Qy 361 AIAGVSAAGVSAAGVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQKILEWEKQ 420
Db 361 AIAGVSAAGVSAAGVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQKILEWEKQ 420
Qy 410 GGQNYFPKGYDSDRYAAYLANNLKFLSELNKELEAERVIAITTOORWNNIGELAGITKLGE 469
Db 410 GGQNYFPKGYDSDRYAAYLANNLKFLSELNKELEAERVIAITTOORWNNIGELAGITKLGE 469
Qy 421 HGKNYFENGVDARVLANLQDNKFLNKLNLKELQAEVIAITTOORWNNIGELAGITKLGE 480
Db 421 HGKNYFENGVDARVLANLQDNKFLNKLNLKELQAEVIAITTOORWNNIGELAGITKLGE 480
Qy 470 RIKSKAYADAFEDGKKEVAGSNITLDATGTTIDISNSNGKKTQALHFTSPLLTAGTESR 529
Db 470 RIKSKAYADAFEDGKKEVAGSNITLDATGTTIDISNSNGKKTQALHFTSPLLTAGTESR 529
Qy 481 KVLGSKAYVDAFEBEGKHLKADKLVLQDSANGIIDVNSNGRAKTKQHILFRTPLTPGTEHR 540
Db 481 KVLGSKAYVDAFEBEGKHLKADKLVLQDSANGIIDVNSNGRAKTKQHILFRTPLTPGTEHR 540
Qy 530 ERLTNGKYSYINKLKFGRVKNQVTDGEASSKLDPSKVIQVQV-----AETEGTDEIGL 582
Db 530 ERLTNGKYSYINKLKFGRVKNQVTDGEASSKLDPSKVIQVQV-----AETEGTDEIGL 582
Qy 541 ERVQTKGYEYITKLININRVDSWKITDGAASSTFDLTNNVQRIEGLNAGNVTKTKETKI 600
Db 541 ERVQTKGYEYITKLININRVDSWKITDGAASSTFDLTNNVQRIEGLNAGNVTKTKETKI 600
Qy 583 IVNAKAGNDDIFVGGQKMNIDGGDGHDRVFSKQDGFNITVDGTSATAGSVTVNRKVAR 642
Db 583 IVNAKAGNDDIFVGGQKMNIDGGDGHDRVFSKQDGFNITVDGTSATAGSVTVNRKVAR 642
Qy 601 VAKLGAGDNDVFGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETQSGSYTVNRFE 659
Db 601 VAKLGAGDNDVFGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETQSGSYTVNRFE 659
Qy 643 RGDYIHEVVKRQETKVGKRTETIOYRDYELRKVGYSQSDNLKSVVEVIGSQFNDVFKG 702
Db 643 RGDYIHEVVKRQETKVGKRTETIOYRDYELRKVGYSQSDNLKSVVEVIGSQFNDVFKG 702
Qy 660 TGKALHEVTSHTALVGNREKIEYR-HSNQNHAGYTTKDTLKAIVEIIGTSHNDIPKG 718
Db 660 TGKALHEVTSHTALVGNREKIEYR-HSNQNHAGYTTKDTLKAIVEIIGTSHNDIPKG 718

```

```

Qy 703 SKENDI FHS GEGD LLDGGAGDRL FCGKGNDRSLGDEGDDLLDGGSGDDVILGGAGNDV 762
Db 719 SKENDAFNGGDDVTIDNGDGNDRFLGKGDDIIDDGGNGDDFDIDGGKGNOLLHGGKGDDI 778
Qy 763 YIFKAGDNDTLVDGTGNDKLAFADANISDIMEIKTEGIIIVKRNHSGGSINIPRWY--- 819
Db 779 FVHRQGDNDIITDSGNDKLSFSDSNLKDLTFKVKHNLVI-TNSRKEKVTIQDWFEA 837
Qy 820 -ITSNLQYQSNKTHKIEQIIGKDGSIYTSDDIKILQDKKGTVITSQELKKLADENK 878
Db 838 DFAKEVRYNATK-DEKIEEIIQNGERITSKQVDDLI--AKNGKKTODELSKVVDNYE 894
Qy 879 SQKLSASDIASSLNKLGVSMALFGTANSVSNALQIPTQPTQGI 922
Db 895 LLKHS-KNVTNSLQKLISSASFTSSNDRNLVAPTSMLDQSL 937

RESULT 3
Q9EV23 PRELIMINARY; PRT; 953 AA.
AC Q9EV23;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Mannheimia glucosida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=85401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH290;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT haemolytica.";
RL J. Bacteriol. 183:1394-1404 (2001).
DR EMBL; AF314522; AAG40306.1; -.
DR InterPro; IPR001753; EnCoA_hydrase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR001995; RtxA.
DR Pfam; PF00353; hemolysinCbind; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102150 MW; D99C36DA595B1624 CRC64;

Query Match 50.7%; Score 2356; DB 2; Length 953;
Best Local Similarity 49.5%; Pred. No. 2.5e-97;
Matches 471; Conservative 182; Mismatches 258; Indels 40; Gaps 15;

Qy 1 MSNINV-IKSN---IQAGLNSI-----KSLGKLYLAIPKD--YDPQKGGTLND 43
Db 5 LTNISTNLKSWLTAKSGLNRTGSLAKAGQSLTKGAKKILYIPKQYQYDTEKNGGLQD 64
Qy 44 FKADELGLLARLAEEPNHTTAKSVYDVTNQFLSLTQTGIAISATKLEKFLQKHSNKL 103
Db 65 LVRAEEUGVEHQEEDNDIAKQTSIGTQNVGLTGERGIVLSAPQDLKLLQK---TKV 121
Qy 104 AKGLDSVENTDKRLGKASNVSLTSLSSFLGTALAGIELDSLTKGDAAPDLAKASIDLIN 163
Db 122 GQAIGSAENLTGKFSNAKTVLSGILGSLVLAGMDLDEALQK-NSNELTLAKAGLELTN 180
Qy 164 EIGNLSQSTQTIKAFSSQAKLGSTTSQAKGFSNIGNKQLNLY-PSKTNLGLBIIITGLL 222
Db 181 SLIENIANSVKTLDAFGDIQNLGSLKQLNVLGSLGDKLGLSFGDKTSILGLDVSGLL 240
Qy 223 SGIAGFALADKNASTGKVAAGFELSNOVIGNTVKAISSVYLAQVAAGLSTTGAAVAL 282

```

```

Db 241 SGATAALVLADKNASTSRKVGAGFELANQVVGNIITKAVSSYILAQRVAAGLSSTGPVAAL 300
Qy 283 ITSIMLAISPLAFMAADKFNHANALDEPAKOFKFGYDGDHLLAAYQRGVGTIEASLT 342
Db 301 IASTVSLAISPLAFAGTADKFNHAKSLESYAERFKLGYDGDHLLAAYQRGVGTIDASVT 360
Qy 343 TISTALGASAGVSAAGVAVGTPIALLVAGVTGLISGLEASQKAMPFESVANRLQKI 402
Db 361 AINTALAAIAGVSAAGVAVGTPIALLVAGVTGLISGLEASQKAMPFESVANRLQKI 420
Qy 403 LEWEKQNGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIATIQORWNNIGELA 462
Db 421 VEVEKNHGNKYFENGVDARYLANLQDNMFKFLNLKELQAERVIATIQOQWNNIGELA 480
Qy 463 GITKLGRIKSGKAYADAPEDGKKVAGSNITIDAKTGIIIDISNCKTKTQALHFTSPLL 522
Db 481 GISRLGKVLGSKAYADAPEDGKKVAGSNITIDAKTGIIIDISNCKTKTQALHFTSPLL 540
Qy 523 TAGTESERLITNGKYSYINKLKFGVRKQWVOTDGEASSKLDKFSKVQIV-----AETE 575
Db 541 TPGTEHRERIQTKYIYITKLNINRVDSWKITDGAASSFTDLTNVQRIEILDNAGNVT 600
Qy 576 GTDEIGLIVNAKAGNDIIFVGOQKMNIDGGDGHDRVYFSKXGGFGNITVDGTSATEAGSY 635
Db 601 KTKETKIVAKLGAGDDNVFVSGTTEIDGEGYDVRVHSR-GNYGALTIDATKETEQGSY 659
Qy 636 TVNRKVARGDIYHEVVKRQETKVGKTTETIQYRDYELRVKVGYSQSTDNLKSVEEVIGSQ 695
Db 660 TVNRFVETGKALHEVTSHTALVGRBEKIEYR-HSNNQHHAGYTTKDTLKAVERIIGTS 718
Qy 696 FNDVFKSGKFNDFHSGEGDLDGAGDDRLFGKGNDRSLGDEGDDLLDGGSGDDVILN 755
Db 719 HNDIFKSGKFNDFHSGEGDLDGAGDDRLFGKGNDRSLGDEGDDLLDGGSGDDVILN 778
Qy 756 GGAGNDVYIPKRGDNDTLVDGTGNDKLAFADANISDIMEIKTEGIIIVKRNHSGGSINI 815
Db 779 GKGDDIIFVHRQGDNDIITDSGNDKLSFSDSNLKDLTFKVKHNLVI-TNSRKEKVTI 837
Qy 816 PRWY-----ITSNLQYQSNKTHKIEQIIGKDGSIYTSDDIKILQDKKGTVITSQELK 871
Db 838 QDWFEADFAKEVRYNATK-DEKIEEIIQNGERITSKQVDDLI--AKNGKKTODELS 894
Qy 872 KLADENKSQKLSASDIASSLNKLGVSMALFGTANSVSNALQIPTQPTQGI 922
Db 895 KVDNVELLKHIS-KNVTNSLQKLISSASFTSSNDRNLVAPTSMLDQSL 944

RESULT 4
Q9ETX2 PRELIMINARY; PRT; 953 AA.
AC Q9ETX2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Mannheimia glucosida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=85401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH498, and PH344;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence diversity and molecular evolution of the leukotoxin (lktA)
RT gene in bovine and ovine strains of Mannheimia (Pasteurella)
RT haemolytica.";
RL J. Bacteriol. 0:0-0 (2001).
DR EMBL; AF314518; AAG40302.1; -.
DR EMBL; AF314517; AAG40301.1; -.
DR InterPro; IPR001753; EnCoA_hydrase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR001995; RtxA.

```



```

Db 481 GISRLGKVLGKAYVDAFEKGHLKADKLVLQDSANGILDVNSGKAKTQHILFRTPLL 540
Qy 523 TAGTESRLTNGKYSYINKLKFRVRNMQVTDGEASSKLDKFSKVIQV-----AETE 575
Db 541 TPGTEHREVRQTKGYEYITKLNINRVDSWKITDGAASSTFDLTNNVQRIGIELDNAGNVT 600
Qy 576 GTDEIGLIVNAKANDDI FVGQKMNIDGGDGHDRVFSKDGFGNITVDGTSATEAGSY 635
Db 601 KTKETKIVAKLGADDDNVFVSGTTEIDGGEYDRVHYSR-GNYGALTIDATKETEQGSY 659
Qy 636 TVNRKVARGDIYHEVVVRQETKVGKRTETIQYRDYELRKVGYGQSTDNLSKVEEVIGSQ 695
Db 660 TVNRVETGKALHEVTSHTALVGNREEKIEYR-HSNNQHAGYTYTKDLKAVEEIIIGTS 718
Qy 696 FNDVFKSGKFNDFPHSGEGDLDLGGAGDDRLFGGKGNDRLSGDEGDDLDGSGDDVLN 755
Db 719 HNDIFKSGKFNDAFNGGDDVDTIDGNDGNRLFGGKGDDIIDGGNGDDFDGKGNDLLH 778
Qy 756 GGAGNDVVI PRKGDNDTLVDGTGNDKLA PADANISDIMIERTKEGIIIVKRNDSHSGSINI 815
Db 779 GKGDDIFVHRQGDNDIITDSGNDKLSFSDSLKDLTTEKVKHNLVI-TNSKKEKVTI 837
Qy 816 PRWY-----ITSNLQYOSNKTDRHKIEQLIGKDGSIYTSQDIDKIQDKDGTVITSQELK 871
Db 838 QDMFREADFAKEVRNYKATK-DEKIEEIIQNGERITSKQVDDLI--AKNGKITQDELS 894
Qy 872 KLADENKSKLSASDIASSINKLVGSMALFGTANVSNNALQIPTQPTQGI 922
Db 895 KVDNYELLKHS-KNVVNSLDKLTISSAFTSSNDSSNRNLVAPTSMLDQSL 944

```

RESULT 6

```

Q9EV26 ID Q9EV26 PRELIMINARY; PRT; 953 AA.
AC Q9EV26;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Mannheimia glucosida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=85401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH240;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Sclander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT haemolytica.";
RL J. Bacteriol. 183:1394-1404 (2001).
DR EMBL; AF14519; AAG40303.1; -.
DR InterPro; IPR001753; EnCoA_hydrtase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RTXa.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102236 MW; AC5874B3B75D1C98 CRC64;

```

```

Query Match 50.5%; Score 2345; DB 2; Length 953;
Best Local Similarity 48.9%; Pred. No. 7.9e-97;
Matches 468; Conservative 186; Mismatches 250; Indels 54; Gaps 16;

Qy 1 MSNTINV-IKSN---IQAGLNST-----KSLGNLYLAIPKD--YDPQKGGLND 43
Db 5 LTNIStNLKSSWLTKSGLNRTGSLAKAGQSLTKGAKKIILYIPKDYQYDTEKGNGLD 64

```

```

Qy 44 FIKAADLGIARLAEPNHTTAKKSVDTVNOFLSLTQTGIAISATKLEKFLQKHSNKL 103
Db 65 LVAAEELGIEVQKEGNDIAKQTSIGTIQWVGLTERGIVLSAQDLKLOK---TKV 121
Qy 104 AKGLDSVENIDRKLKGNVSLTSLSFGLTAGIELDSLIKGDAAAPDAKASIDLIN 163
Db 122 GOAIGSAENLTGKFSNAKTVLSIGISLVLAGMDLDEALQK-NSNELTLAKAGLELTN 180
Qy 164 EITGNLSQSTOTTIEAESSOLAKLGSITISQAKGSNTGNKLQNLN-PSKTNLGLIETGL 222
Db 181 SLIENTANSVKTLDAFGDQINQLGSKLVNQYKLSIGDLKLGSLGDFKSLGDLVVSGLL 240
Qy 223 SGISAGFALADKNASTGKVAAGFELSNOVIGNVTKAISYVLAQRAVAGLSTTGAVAL 282
Db 241 SGATAALVLADKNASTSRKVAGFELANQVGNITKAVSYIILAQRAVAGLSTGPAAL 300
Qy 283 ITSSIMLAISPLAPMAADKFNHANLDEFKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
Db 301 IASTVSLAISPLAFAGIADKFNHAKLSLEYAERFKLGYDGDNLLAEYQRGVGTIDASVT 360
Qy 343 TISTALGAVSAGVSAAGVAVCTPIALLVAGVTGLISGLEASKOAMPESVANRLOGKI 402
Db 361 AINTALAAIAGGVSAAGSVIASPALLVSGITGVTITLOYSKQAMFHVANKIHNI 420
Qy 403 LEWEKONGGONYFPDKGVDSRYAAYLANNLKFLSELNKELEAEARVIAITQORWNNIGELA 462
Db 421 VEWEKNNHKNYFENGYDARYLANLQDNWKFLLNLNKELOAERVIAITQOOWNNIGDLA 480
Qy 463 GITKLGERIKSKAYADAFEDGKKVAGSNITLDAKTGIIIDISNKGKTOALHFTSPLL 522
Db 481 GISRLGKVLGKAYVDAFEKGHLKADKLVLQDSANGIIDVNSGKAKTQHILFRTPLL 540
Qy 523 TAGTESRLTNGKYSYINKLKFRVRNMQVTDGEASSKLDKFSKVIQV-----AETE 575
Db 541 TPGTEHREVRQTKGYEYITKLNINRVDSWKITDGAASSTFDLTNNVQRIGIELDNAGNVT 600
Qy 576 GTDEIGLIVNAKANDDI FVGQKMNIDGGDGHDRVFSKDGFGNITVDGTSATEAGSY 635
Db 601 KTKETKIVAKLGADDDNVFVSGTTEIDGGEYDRVHYSR-GNYGALTIDATKETEGRSY 659
Qy 636 TVNRKVARGDIYHEVVVRQETKVGKRTETIQYRDYELRKVGYGQSTDNLSKVEEVIGSQ 695
Db 660 TVNRVETGKALHEVTSHTALVGNREEKIEYR-HSNNQHAGYTYTKDLKAVEEIIIGTS 718
Qy 696 FNDVFKSGKFNDFPHSGEGDLDLGGAGDDRLFGGKGNDRLSGDEGDDLDGSGDDVLN 755
Db 719 HNDIFKSGKFNDAFNGGDDVDTIDGNDGNRLFGGKGDDIIDGGNGDDFDGKGNDLLH 778
Qy 756 GGAGNDVVI PRKGDNDTLVDGTGNDKLA PADANISDIMIERTKEGIIIVKRNDSHSGSINI 815
Db 779 GKGDDIFVHRQGDNDIITDSGNDKLSFSDSLKDLTTEKVKHNLVI-TNSRKEKVTI 837
Qy 816 PRWY-----ITSNLQYOSNKTDRHKIEQLIGKDGSIYTSQDIDKILQ-----DKKD-GTV 864
Db 838 QDMFREADFAKEVRNYKATK-DEKIEEIIQNGERITSKQVDELBIEKGKIDKSDLSQV 896
Qy 865 ITSQELKKLADENKSKLSASDIASSINKLVGSMALFGTANVSNNALQIPTQPTQGI 922
Db 897 VDNYTELLK-----HSKNVNTSLDKLISSAFTSSNDSSNRNLVAPTSMLDQSL 944

```

RESULT 7

```

Q9EV32 ID Q9EV32 PRELIMINARY; PRT; 953 AA.
AC Q9EV32;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.

```



```

Qy 283 ITSSIMLAISPLAFMNAADKFNHANALDERAKQFRKFGYDGDHLLAAYQGVGTIEASLT 342
Db 301 IASTVSLAISPLAFAGIADKFNHAKSLUESAERFKLQYDGDNLLAAYQGVGTIDASVT 360
Qy 343 TISTALGAVSAGVSAAGVAVGTPIALLVAGVTGLISGLEASKQAFESVANRLQSKI 402
Db 361 AINTALAAAGVSAAGVAVGTPIALLVAGVTGLISGLEASKQAFESVANRLQSKI 420
Qy 403 LEWEKQNGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVAITQORWNNIGELA 462
Db 421 VEMEKNPKGNYFENGVDARYLANLQDNMFLNLNKLQAEARVAITQOQWNNIGELA 480
Qy 463 GITKLERIKSGKAYADAFEDGKVEAGSNITLDAKTGIIDISNCKKQALHFTSPLL 522
Db 481 GISRLGKSVLGGKAYDAFEGQHLKADKLQVLDGAGKIIDVSNTEGAKTQHLFRTPLL 540
Qy 523 TAGTESERLTNGKYSYINKLKFRVRKNQWVTDGEASSKLDKFSKVIQV-----AETE 575
Db 541 TPGTEKERVQTKGYEYITKLHNRVDSWKITDGAASSTFDLTNNVQRIEGLDNAGNVT 600
Qy 576 GTDEIGLIVNAKAGNDIDFVGGQKRNIDGGDHRVYFKDGGFGNITVDTGSATEAGSY 635
Db 601 KTKETKIIAKLGGDDNVFVSGGTTIDGGEYDRVHYSR-GNYGALTIDATKETEQQSY 659
Qy 636 TVNRKVARGDYIHEVVKRQETKTKRTETQYRDYELRVKVGYGQSTDNLSKVEEVIGSQ 695
Db 660 TVNRFVSGKALHEVTSHTALVGNREEKIEYR-HSNNOHHAGYITKDTLKAVEEIGTS 718
Qy 696 FNDVFKGSKENDIFHSGEGDLDLGGAGDRLGCGKGNDRLSGDEGDDLDGSGDDVLN 755
Db 719 HNDIFKSGKENDAFNGGDDVTIDGNDGRLFGKGDDIIDGCGDDFDGKGNDLLH 778
Qy 756 GGAGNDVYIFRKGNDTLVDGTGNDKLA PADANISDIMIERTKEGIIIVKRNDSHGSINI 815
Db 779 GKGDDIFVRHQGDNDSTISEGNDKLSFSDSNLKDLTFEKVNHLVI-TNTKQEKVTI 837
Qy 816 PRWY----ITSNLQNSKTDHKIEQLIGKGSYVITSDQDKLQDKDGTWITSDELK 871
Db 838 QNWREAEFAKTIQNYVATR-DDKIEEIQNGRITSKQVDELIE--KNGKIAQSELT 894
Qy 872 KLADENKSQKLSASDIASSLNKLVGSMALFCTANSVSNALQIPTQGTGILAPSV 927
Db 895 KVDNYQLLKYS-RDASNSLDKLSSASAFSSND-SRNVL---ASPT-SMLDPSL 944

RESULT 9
Q9EV28 ID Q9EV28 PRELIMINARY; PRT; 953 AA.
AC Q9EV28 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leukotoxin (Lkta).
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH292, and PH296;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA) Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella) haemolytica.";
RT J. Bacteriol. 183:1394-1404(2001).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=PH296;
RA Davies R.L., Campbell S., Whittam T.S.;
RT "Mosaic structure and molecular evolution of the leukotoxin operon

```

```

RT (lktCABD) of Mannheimia (pasteurella) haemolytica, Mannheimia
RT Glucosida and Pasteurella trehalosi.";
RL J. Bacteriol. 0-0-0(2001).
DR EMBL; AF314515; AAG40299.1; -.
DR EMBL; AF414141; AAL13281.1; -.
DR InterPro; IPR001753; EnCoA_hydrtase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; Rtx, 1.
DR PRINTS; PR00313; CABDNRPRT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL COA HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102218 MW; 04AB1715B819E571 CRC64;

Query Match 50.2%; Score 2332.5; DB 2; Length 953;
Best Local Similarity 49.8%; Pred. No. 2.8e-96;
Matches 474; Conservative 180; Mismatches 257; Indels 45; Gaps 18;

Qy 1 MSNINV-IKSN---IQAGLNST-----KSGLNLYLAIPKD--YDPQKGGTLND 43
Db 5 LTNISTNLKSWTAKSGLNRTQSLAKAGOSLKTAKKILYIPKDYQYDTDKGNLQD 64
Qy 44 FIKAADELGIARLAERPNTHTAKKSVDTVNQFLSLTQTGTGIAISATKLEKLEKQKSTNKL 103
Db 65 LVKAAEELGIEVQKESSNDIAKAQTSGLTTHNVGLTERGIVLSAPQLDKLQK---TKV 121
Qy 104 AKGLDSVENIDRKLKGNVSLSTLSSFLGTALAGIELDSLKKGAADPAALAKASIDLIN 163
Db 122 GQAIGSTENITKGFNAKTVLSGILSGVSLGMDLDEAL-QNNSNELTLAKAGLELTN 180
Qy 164 EIIGNLSOSTQTTEAFSSOLAKLGSTISQAKGFSNIGKNLQNLN-FSKTNLGLIEITGLL 222
Db 181 SLIENTANSVKTLDAFDQINQLGSKLVNKGSLSSGEKLGSLGDPKTSGLGDIVSGLL 240
Qy 223 SGISAGFALADKNASTGKVAAGFELSNQVIGNVTKAISYVLAQRAAGLSTTGVAAL 282
Db 241 SGATAALVLADKNASTRKVGAGFELANQVGNITKAVSSYILAQRAAGLSTGTPVAAL 300
Qy 283 ITSSIMLAISPLAFMNAADKFNHANALDERAKQFRKFGYDGDHLLAAYQGVGTIEASLT 342
Db 301 IASTVSLAISPLAFAGIADKFNHAKSLUESAERFKLQYDGDNLLAAYQGVGTIDASVT 360
Qy 343 TISTALGAVSAGVSAAGVAVGTPIALLVAGVTGLISGLEASKQAFESVANRLQSKI 402
Db 361 AINTALAAAGVSAAGVAVGTPIALLVAGVTGLISGLEASKQAFESVANRLQSKI 420
Qy 403 LEWEKQNGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVAITQORWNNIGELA 462
Db 421 VEMEKNPKGNYFENGVDARYLANLQDNMFLNLNKLQAEARVAITQOQWNNIGELA 480
Qy 463 GITKLERIKSGKAYADAFEDGKVEAGSNITLDAKTGIIDISNCKKQALHFTSPLL 522
Db 481 GISRLGKSVLGGKAYDAFEGQHLKADKLQVLDGAGKIIDVSNTEGAKTQHLFRTPLL 540
Qy 523 TAGTESERLTNGKYSYINKLKFRVRKNQWVTDGEASSKLDKFSKVIQV-----AETE 575
Db 541 TPGTEKERVQTKGYEYITKLHNRVDSWKITDGAASSTFDLTNNVQRIEGLDNAGNVT 600
Qy 576 GTDEIGLIVNAKAGNDIDFVGGQKRNIDGGDHRVYFKDGGFGNITVDTGSATEAGSY 635
Db 601 KTKETKIIAKLGGDDNVFVSGGTTIDGGEYDRVHYSR-GNYGALTIDATKETEQQSY 659
Qy 636 TVNRKVARGDYIHEVVKRQETKTKRTETQYRDYELRVKVGYGQSTDNLSKVEEVIGSQ 695
Db 660 TVNRFVSGKALHEVTSHTALVGNREEKIEYR-HSNNOHHAGYITKDTLKAVEEIGTS 718
Qy 696 FNDVFKGSKENDIFHSGEGDLDLGGAGDRLGCGKGNDRLSGDEGDDLDGSGDDVLN 755
Db 719 HNDIFKSGKENDAFNGGDDVTIDGNDGRLFGKGDDIIDGCGDDFDGKGNDLLH 778
Qy 756 GGAGNDVYIFRKGNDTLVDGTGNDKLA PADANISDIMIERTKEGIIIVKRNDSHGSINI 815

```

```
Db 779 GKGDDIFVHRQGDGNDISITSEGNKLSFSNLSKDLTFEKNVHHLVI--TNTKQEKVTI 837
Qy 816 PRWY-----ITSNLQYQNKTDHKEQLIGKDGSVITSQDKLQDKKQDGTVTTSQELK 871
Db 838 QNWFEAEFAKTIQNYVATR-DDKTEIIGQNGERITSQVDELTE--KNGKIAQSELT 894
Qy 872 KLADENKQKLSASDIASLKNLVCMSMALFGTANSVSSNALOPTQPTQGLAPSV 927
Db 895 KVDVNYQLLKYS--RDASNLDKLITSSAFTSSND--SRNLV---ASPT--SMLDPSL 944

RESULT 10
Q3ETG5 PRELIMINARY; PRT; 953 AA.
AC Q9ETG5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH284, PH388, and PH8;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence diversity and molecular evolution of the leukotoxin (lktA)
RT gene in bovine and ovine strains of Mannheimia (Pasteurella)
RT haemolytica";
RL J. Bacteriol. 0:0-0(2001).
DR EMBL; AF314507; AAG40291.1; -.
DR EMBL; AF314504; AAG40288.1; -.
DR EMBL; AF314506; AAG40290.1; -.
DR InterPro; IPR001753; EnCoA_hydrtse.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabinid; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR01488; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ SEQUENCE 953 AA; 101997 MW; D593DeA577C3AD59 CRC64;

Query Match 50.2%; Score 2332; DB 2; Length 953;
Best Local Similarity 50.1%; Pred. No. 3e-96;
Matches 465; Conservative 175; Mismatches 263; Indels 26; Gaps 13;

Qy 8 KSNIOAGLNTSKGLKNIYLAIPKD--YDPQKGTFLNDFIKAADELGTARLAERPHET 65
Db 28 QSLTQAG--SSLTKGAKIILYIPQNYQYDTQNGLQDLVAAAEELGIEVQREERNNTAT 86
Qy 66 AKKSVDTVNOFLSTQGTATISATKLEFLQKHSTNKLAKGLDSVENIDRKLGSANVLS 125
Db 87 AQTSIGTTQTAIGLTERGIVLSAPQDKLQK---TKAQLGSAESIVQANRAKTIVLS 143
Qy 126 TLSSFLGTALAGIELDLSLKGDAAAPDALAKASIDLINEIINLSQSTQTTEAFSSQAK 185
Db 144 QIOSILGSLVAGMDLDEAL-QNNNSQHALAKAGLELTNSLIENIANSVKTLDFEGEQISQ 202
Qy 186 LGSTISOAKGFSNIGNKQNLN-NFSKTNLIGLEIITGLLSGISAGFALADKNASTGKVA 244
Db 203 FGSKLQNTKGLTGLDKLNGGLDKAGLGLDVTISGLLSGATAALVLADKNASTAKKYGA 262
Qy 245 GFELSNQVIGNVTKAISSVLAQRAAGLSTTGAVAALITSSILAIPLAFMAADKFN 304
Db 263 GFELANQVGNITKAVSSITLQRAAGLSSTGPVAAIASTVLSAISPLAFAGIADKFN 322
Qy 305 HANALDEFAKQRFKFGYGDHLLAEYQRGVGTIEASLTITSTALGAVSAGVSAAAVGS 364
Db 323 HAKSLESYAERPKLGYDGDNLAEYQRGVTGTIDASVTAINALAAIAGVSAAGSVI 382
```

```
Qy 365 GTPIALVAGVTGLISGILEASKQAMFESVANRLQGGKI LEWEKONGGONYFDKGYDSRYA 424
Db 383 ASPIALVSGITGVISTILQISQKAMFEHANKIHNKIWEKKNHGKNYFENGYDARYL 442
Qy 425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNNIGELAGITKLGRIKSGKAYADAFEDG 484
Db 443 ANLQDNMKFLNLNKLQAEVIAITQQQWDDNNIGDLAGISRLGEKVLGKAYADAFEEG 502
Qy 485 KKVFAAGNITLDKAGTIIIDISNSNGKQTQALHFTSPLLTAGTBSRRLTNGKYIYNKLK 544
Db 503 KHIKADKLVLQDSANGIIDVSNCGKAKTQHILFRTPLLTGCTEHRERVQTGKYEYITKLN 562
Qy 545 FGRVKWQVTDGEASSKLDPSKVLQV-----AETEGTDEIGLIYNAKAGNDDIFVGQ 597
Db 563 INRVDSWKITDGAASSTFDLTNVVQIRIGIELDNNAGNTKTKETKIIAKLGGDDNVFVGS 622
Qy 598 GKWNIDGGDGHDRVYFSGDKGFGNITVDGTSATSAGSYTVNRRKVARGDIYEVVVKRQETK 657
Db 623 GTTEIDGEGVDRVHYSR-GNYGALTIDATKETEGSGSYTVNRFVETGKALHEVTSHTAL 681
Qy 658 VGKGTETIQRYDYELRKVGYGYQSTDMNLSKVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
Db 682 VGNREEKIEYR-HSNNOHHAGYTYKDTLKAVEEIIGTSHNDIFPKSGFNDAFNGDGYDT 740
Qy 718 LDGGAGDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDNNTLYDG 777
Db 741 IDGNDGNRLFGGKGDDILDGGNGDDIFDGGKGNLLHGGKGGDIFVHRKGGDNDIITDS 800
Qy 778 TGNDKLAFAFADANISDIMIERKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
Db 801 DGNDKLSFSDSNLKDITFEKVKNLVI-TNSKKEKVTIQDWFREADFAKEVPNYKATK-D 858
Qy 834 HKIEQLIGKDGSVITSQDKLQDKKQDGTVTTSQELKKLADENKSKLASDITASSLNK 893
Db 859 EKIEEIIQNGERITSQVDDLI--AKNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 915
Qy 894 LVGSMALFGTANSVSSNALOPTQPTQGI 922
Db 916 LISSVSAFTSSNDSNRNVLVAPTSMLDQSL 944

RESULT 11
Q9EV31 PRELIMINARY; PRT; 953 AA.
AC Q9EV31;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH588;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT haemolytica";
RL J. Bacteriol. 183:1394-1404(2001).
DR EMBL; AF314510; AAG40294.1; -.
DR InterPro; IPR001753; EnCoA_hydrtse.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabinid; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ SEQUENCE 953 AA; 102160 MW; A189BF80754A7907 CRC64;
```

| | | |
|-----------------------|---|--|
| Query Match | | 50.1%; Score 2329.5; DB 2; Length 953; |
| Best Local Similarity | | 50.1%; Pred. No. 3:9e-96; |
| Matches | | 469; Conservative 178; Mismatches 259; Indels 31; Gaps 16; |
| Qy | 5 NVIKSNIOAGLNSKSGLNLYLAIPKD--YDPQKGGTFLNDFIKAADEBELGARLAERPNNH | 62 |
| Db | 25 NAGOSLAKAG-QSLKTGAKKIILYIPKDYDTEKNGQLDLVKAAGELGIEVQKEGND | 83 |
| Qy | 63 TETAKKSVDTVQFLSTQTGTGIAISATKLEKFLQKHSTNKLAKGLSDSVENIDRKLGHKASN | 122 |
| Db | 84 IAKAQTSLGTQNVGLTERGIVLSAPQLDKLQK--TKVQAIGSAENLTGKFSNAKT | 140 |
| Qy | 123 VLSTLSSFLGTALAGIELDSLIIKGDAPDALAKASIDLINELIENLSQSQTTEARSSQ | 182 |
| Db | 141 VLSGIQSLTSLVAGMDLDEALQK-NSNELTLAKAGUELTNSLIENANSVKTLTADFGDQ | 199 |
| Qy | 183 LAKLGSTISQAKGFSNIGKNLQNLN-FSKTNLGLIEITLGLSGISAGFALADKNASTGCK | 241 |
| Db | 200 INQLGSKLQNVKGLSSGLDKLGLSGDPKTSGLDLDVSGLLSGATAALVLADKNASTRK | 259 |
| Qy | 242 VAAGFELSNQVGNVNTKAISSVYLAQRAAGLSTTGAAVALITSSIMLAISPLAFMAAD | 301 |
| Db | 260 VGAGFELANQVGNITRAVSSYILAQRAAGLSTGTPVAALIASTVSLAISPLAFAGIAD | 319 |
| Qy | 302 KFNHANALDEFAPKFRFGYGDHLLAAYQRCVGTIEASLTITISALGANSAGVSAAGV | 361 |
| Db | 320 KFNHAKSLSEYAEFRFKGLYGDHLLAAYQRCVGTIEASLTITISALGANSAGVSAAGV | 379 |
| Qy | 362 SAVGTPTALLVAGVTGLISGLEASQAMPFESVANRLQKLEWEKONGGQNYFDKGYDS | 421 |
| Db | 380 SVIASPALLVSGITGVISTLIQVSKQAMPFHVANKIHNKIVWEKONGGKQNYFENGIDA | 439 |
| Qy | 422 RYAAAYLANNKFLSELNKELEAERVIAITQOQWNNNIGELAGITKLERIKSGRAYADAF | 481 |
| Db | 440 RYLANLQDNMFKFLNLNKLQELQAERVIAITQOQWNNNIGELAGISRLGEKVLGKAYVDAF | 499 |
| Qy | 482 EDGKKVAGSNITLDKAGTIIDISNGKKTQALHFTSPLITAGTESRERLTNGKYSYIN | 541 |
| Db | 500 BEGQHLKADKLVDLSAKGIIDVNTGEBKTQHILFRTPLLTPGTGKERVQTKGYEIT | 559 |
| Qy | 542 KLKFRVQNVQVTDGEASSKLDPSKVIQORVA-----ETEGTDEIGLIVNAKAGNDDIF | 594 |
| Db | 560 KLHINRVDSWQIKDGAASSTFDLTNVQIRIGVELDHAENVIKETKIVATLGGDDNVF | 619 |
| Qy | 595 VQGGMNIDGGHDRVFPYSKDGFGNITVDGTSATEAGSYTVNRKVARGDYIHEVVQRQ | 654 |
| Db | 620 VSGTTEIDGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVESGKALHEVTS | 678 |
| Qy | 655 ETKVGRKTETIOYRDYELRKVGYQSTDNILKSVEVIGSQFNDVFKGSKENDIFHSQEG | 714 |
| Db | 679 TALVGNREEKTEYR-HSNQHHAGYTYKDTLKAVEEIIIGTSHNDIFKSGKFNDAFGDQ | 737 |
| Qy | 715 DDLDDGAGDRLFGKGNDRSLGDEGDDLDGSGDDVLNGAGNDDVYIPRKGDGNDTL | 774 |
| Db | 738 VDTIDGNDGRLFGKGGDDIIDGNGDDFDGKGKNDLHGGKDDIFVHRQSGNDISI | 797 |
| Qy | 775 YDGTGNDKLPADANISDMITERTKEGIIIVKRNDSHSGSINIPRWY----ITSNIQNYQSN | 830 |
| Db | 798 TESGNDKLSFSDNLKDLTEKYNHHLVI-TNTQKEKVTIQNFREAEFAKTIQNYVAT | 856 |
| Qy | 831 KTDHKEIOLIGKSGSYTSDQIKLQDKDGTVITSQELKLADEKNSQKLSASDASS | 890 |
| Db | 857 R-DDKIEEIIQONGERITTSKVDELIE--KNGKIAQSELTKVVDNYQLLKYS-RDASNS | 912 |
| Qy | 891 LNKLVGSMALFGTANSVSSNALQPIPTOPTGILAPSV | 927 |
| Db | 913 LDKLISSASATSSND-SRNVL---ASPT-SWLDPSL | 944 |
| RESULT 12 | | |
| Q9EV27 | | |
| ID | Q9EV27 | PRELIMINARY; PRT; 953 AA. |

| | | |
|--|--|-----|
| AC | Q9EV27; | |
| DT | 01-MAR-2001 (TrEMBLrel. 16, Created) | |
| DT | 01-MAR-2001 (TrEMBLrel. 16, Last sequence update) | |
| DT | 01-MAR-2003 (TrEMBLrel. 23, Last annotation update) | |
| DE | Leukotoxin. | |
| GN | LkTA. | |
| OS | Pasteurella haemolytica. | |
| OC | Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; | |
| OC | Pasteurellaceae; Mannheimia. | |
| OX | NCBI_TaxID=75985; | |
| RN | [1] | |
| RP | SEQUENCE FROM N.A. | |
| RC | STRAIN=PH392; | |
| EX | MEDLINE=21101823; PubMed=11157953; | |
| RA | Davies R.B., Whittam T.S., Selander R.K.; | |
| RT | "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA) | |
| RT | Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella) | |
| RT | haemolytica."; | |
| RL | J. Bacteriol. 183:1394-1404 (2001). | |
| DR | EMBL: AF314516; AAC40300.1; - | |
| DR | InterPro; IPR001753; EnCoA_hydrasee. | |
| DR | InterPro; IPR001343; Hemylyn_Ca_bind. | |
| DR | InterPro; IPR003995; RtxA. | |
| DR | Pfam; PF00353; hemolysinCabin; 5. | |
| DR | Pfam; PF02382; Rtx; 1. | |
| DR | PRINTS; PR00313; CABNDNGRPT. | |
| DR | PRINTS; PR01488; RYTOXINA. | |
| DR | PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1. | |
| SQ | SEQUENCE 953 AA; 102230 MW; 2B686808EB370090 CRC64; | |
| Query Match | | |
| Best Local Similarity 50.1%; Score 2329.5; DB 2; Length 953; | | |
| Matches 474; Conservative 180; Mismatches 257; Indels 45; Gaps 18; | | |
| Qy | 1 MSNINV-IKSN--IQAGLNST-----KSLKLNLYLAIPKD--YDPQKGGTIND | 43 |
| Db | 5 LTNIWNLKSSWLTAKSGLNRTGSLAKAGQSLKTKGAKKIILYIPKDYQYDQDCKNGQLQD | 64 |
| Qy | 44 FIKAADEBELGARLAERPNNHETAKKSVDTVQFLSTQTGTGIAISATKLEKFLQKHSTNKL | 103 |
| Db | 65 LVKAAEELGIEVQKEESNDIAKQTSGLTTHNVGLTERGIVLSAPQLDKLQK--TKV | 121 |
| Qy | 104 AKGLSDSVENIDRKLKGNASVNLSTLSSFLGTALAGIELDSLIIKGDAPDALAKASIDLIN | 163 |
| Db | 122 GQAIGSTENTIKGFSNAKIVLSGIQSLGSLVAGMDLDEAL-QNNSNELTLAKAGLELTN | 180 |
| Qy | 164 EITGNLSQSQTTEAFSSQAKLGSTTISQAKGFSNIGKNLQNLN-FSKTNLGLIEITGLL | 222 |
| Db | 181 SLIENIANSVKTLDAFGDQINQLGSKLQNVKGLSSLGEKLGSGFKTSGLGLDIVSGLL | 240 |
| Qy | 223 SGISAGPALADKNASTGKKVAAGFELSNOVIGNVTKAISSVLAQRAAGLSTTGAAVAL | 282 |
| Db | 241 SGATAALVLADKNASTRKVKVAGFELANQVGNITKAVSSYILAQRAAGLSTGTPVAAL | 300 |
| Qy | 283 ITSIMLAISPLAFMAADKFNHANALDEFAPKFRFGYGDHLLAAYQRCVGTIEASLT | 342 |
| Db | 301 IASTVSLAISPLGFPAGTADKFNHAKSLSEYAEFRFKGLYGDHLLAAYQRCVGTIEASVT | 360 |
| Qy | 343 TISTALGAVSAGVSAAGVSAVGTPTALLVAGVTGLISGLEASQAMPFESVANRLQKI | 402 |
| Db | 361 AINTALAAIAGGVSAAGSVASPIALLVSGITGVISTLIQVSKQAMPFHVANKIHNI | 420 |
| Qy | 403 LEWEKONGGQNYFDKGYDSRYAAYLANNKFLSELNKELEAERVIAITQOQWNNNIGELA | 462 |
| Db | 421 VEWEKNNPGNYFENGIDARYLANLQDNMFKFLNLNKLQELQAERVIAITQOQWNNNIGELA | 480 |
| Qy | 463 GITKLERIKSGRAYADAFEDGKKVVEAGSNITLDKAGTIIDISNGKKTQALHFTSPLL | 522 |
| Db | 481 GISRLGKVLGSKAYVDFAEQQHKLKADKLVDLSAKGIIDVNTGEBKTQHILFRTPLL | 540 |
| Qy | 523 TAGTESRERLTNGKYSYINKLKFGRVKNQWVTDGEASSKLDPSKVIQORV-----AETE | 575 |
| Db | 541 TPGTEKEERVQTKGYEITKLIHNRVDSWKITDGAASSTFDLTNVQVQIRIGIELDNAGNVT | 600 |


```
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Pasteurella trehalosi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=47735;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH68;
RX MEDLINE=21101823; PubMed=11157953;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence Diversity and Molecular Evolution of the Leukotoxin (lkta)
RT Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RL haemolytica.";
RL J. Bacteriol. 183:1394-1404 (2001).
DR EMBL; AF314526; AAG40310.1; -.
DR InterPro; IPR001753; EnCoA_hydrase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003955; RtxA.
DR Pfam; PF00353; hemolysinCbind; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ SEQUENCE 955 AA; 102069 MW; FDCFZE99B8DC394F CRC64;

Query Match 49.2%; Score 2286; DB 2; Length 955;
Best Local Similarity 49.0%; Pred. No. 3.4e-94;
Matches 466; Conservative 177; Mismatches 264; Indels 44; Gaps 17;

Qy 5 NVIKSNTQA---GLN-----STKSLGNLYLAIPKD--YDPQKGGTLNDFTKAA 48
Db 12 NGIRSTLTATRGGLNAGQSLTQAGTLKNGAKKIILYIPKDYKDSGSGNGLQDLVKAA 71
Qy 49 DELGIARLAEPNHTETAKKSDVTNQFLSLTQTGTIAISATKLEFLQKHSTNKLAKGLD 108
Db 72 EELGIEVQKEEGNDIAKAQTSGLTQNVGLGTERGIVLSAPQLDKLQK---NKVGQALG 128
Qy 109 SVENIDRKLKASNVLSLSSFLGTALAGIELDSLKKGDPAADALAKASIDLINELIGN 168
Db 129 SSESIAQNFQSAKTVLSGVQSVILSGVLGMDLDEAL-QNESDQLTLAKAGLELNSLIEN 187
Qy 169 LSQSTQTEAPSSQAKLGSTISQAKGSFNIGNKQNL-NFSKTNVLGEIITGLLSGISA 227
Db 188 IANSVQTLDAFSEQISQFSGSKLVNQVKGILGALGDKLKNIGGLDKAGLGLDVISGLSGATA 247
Qy 228 GFALADKNASTGKKAAGFELSNOVIGNVTKAISVYLAORVAAGLSTTGGAVALITSSI 287
Db 248 ALVLADKADASTAKKVGAGFELANQVGNITKAVSYILAORVAAGLSTTGGAVALIASTV 307
Qy 288 MLATSPALFMAADKFNHANALDEFAKFRFGYDGDHLLAEYQGVGTIEASLTITSTA 347
Db 308 ALATSPISFAGIAKDFRAKSLLENYAERFKLGEGLSLLAEYOHGTIDASVTALNTA 367
Qy 348 LGAVSAGVSAAGVAVGTPTIALIVAGVTGLISGILEASKQAMPESVANRLOGKILEWEK 407
Db 368 LAATAGGVSAAGVSAAPVALLVSGITGVISTILQYSKQAMPFHVANKHNKIVWEK 427
Qy 408 QNGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQORDNNIGELAGITKL 467
Db 428 NNGGKNYPFENGYDARYLANLQDNMKFLNLNKLQEAERVIAITQQQWDSNIGLAGISRL 487
Qy 468 GERIKSGKAVADAFEDGKVKVAGSNI TLDAKTGIIDISNSNGKITQALHFTSPLLTAGTE 527
Db 488 GEKVLGKAVYDAFEEGHLKADKLVLDSAKGIIDVSNTEAKTQHLFPTLLTGTTE 547
Qy 528 SREBLTNGKYSYINKLKFGRVKNHQTVDGEASSKLDPSKVTQVRA-----ETEGTDEI 580
Db 548 KRERVQTKGYEYTKLHINRVDSNQIKDGAASSTFDLTNVVQRTGVLDHAENVIKTKET 607
```

```
Qy 581 GLIYNKAGNDDIFVGQKKNIDGGDGHDRVYFSKDGFGNITVDGTSATEAGSYTVNRK 640
Db 608 KIVATLGGDDNVFVSGTTEIDGEGYDRVHYSR-GNYGALTIDATKETEQSSYTVNRF 666
Qy 641 VARGDIYHEVVKRQETKRTETIOYRDYELARKVGYGOSTNLKSVVEVIGSQFNDVF 700
Db 667 VESGKALHEVTSTHTALVGNREEKIEYR-HSNQHAGYYTKDTLKAVEEIIIGTSHNDF 725
Qy 701 KGSKENDDIFHSGRGDDLLDGGAGDDRLFGGKGNDRLSGDGDDLLDGGSGDDVLNGGAGN 760
Db 726 KGSKFNDAFNGGQVDTIDGNGNDRLFGGKGDDIDGNGGDDFIDGGKGNLLHGGKGD 785
Qy 761 DVYIFRKGDNGLTYDGTGNDKLAFADANISDIMIERKEGIIVKRNDSHSGSINIPRWY- 819
Db 786 DIFVHRQGDGNDSTITESGNDKLSFSDSNLKDITFEKVNHLVI-TNTQKEKVTIQWFR 844
Qy 820 ---ITSNLQYQSNKTDHKEIQIKDGSVITSQDKILODKKDGTVITVSQELKLADE 876
Db 845 EAEFAKTIQYVATR-DDKIEEIIQNGERITSQVDELIE--KGNGKIAQSELTKVVDN 901
Qy 877 NKSQKLSASDIASSLANKVCSMALFGTANSVSSNALQPIQTQGTILLAPSV 927
Db 902 YQLLKYS-RDASNDLKLISASFTSSND-SRNL---ASPT-SMLDPSL 946

RESULT 18
Q9EUD4
ID Q9EUD4 PRELIMINARY; PRT; 955 AA.
AC Q9EUD4;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 23, Last annotation update)
DE Leukotoxin.
GN LKTA.
OS Pasteurella trehalosi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=47735;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH254, PH246, and PH252;
RA Davies R.L., Whittam T.S., Selander R.K.;
RT "Sequence diversity and molecular evolution of the leukotoxin (lkta)
RT gene in bovine and ovine strains of Mannheimia (Pasteurella)
RT haemolytica";
RL J. Bacteriol. 0:0-0 (2001).
DR EMBL; AF314525; AAG40309.1; -.
DR EMBL; AF314523; AAG40307.1; -.
DR EMBL; AF314524; AAG40308.1; -.
DR InterPro; IPR001753; EnCoA_hydrase.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003955; RtxA.
DR Pfam; PF00353; hemolysinCbind; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ SEQUENCE 955 AA; 102127 MW; 074B2E4ADF5F7AA CRC64;

Query Match 48.9%; Score 2274; DB 2; Length 955;
Best Local Similarity 48.7%; Pred. No. 1.2e-93;
Matches 463; Conservative 180; Mismatches 264; Indels 44; Gaps 17;

Qy 5 NVIKSNTQA---GLN-----STKSLGNLYLAIPKD--YDPQKGGTLNDFTKAA 48
Db 12 NGIRSTLTATRGGLNAGQSLTQAGTLKNGAKKIILYIPKDYKDSGSGNGLQDLVKAA 71
Qy 49 DELGIARLAEPNHTETAKKSDVTNQFLSLTQTGTIAISATKLEFLQKHSTNKLAKGLD 108
Db 72 EELGIEVQKEEGNDIAKAQTSGLTQNVGLGTERGIVLSAPQLDKLQK---NKVGQALG 128
Qy 109 SVENIDRKLKASNVLSLSSFLGTALAGIELDSLKKGDPAADALAKASIDLINELIGN 168
```



```
Db 129 SSESIAQNFQSAQTVLSCVQSIILGVLAGMDLDEAL-QNESDQLTLAKAGLELTNSLIEN 187
Qy 169 LSQSTQTIAPSSQIAKLGSTISQAKGFSNIGKQML-NFSKTNLGLLEITGLLSGISA 227
Db 188 IANSVQTLDAFSEQISQFSGKLQNVKGLGALGDKLNKIGGLDKAGLGLDVISGLSGATA 247
Qy 228 GFALADKNASGKVAAGFELSNOVIGNVTKAISYVLAQRAAGLSTTGAVALIITSSI 287
Db 248 ALVLADKDASTAKVAGAGFELANVGNITKAVSYILAQRAAGLSTGTPVALIASTV 307
Qy 288 MLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQGVGTIEASITTTSTA 347
Db 308 ALAISPLSFAGIADKFRDRAKSLNVAERFKLGYEGDSLLAEYQHGTTIDASVTAINTA 367
Qy 348 LGAVSAGVSAAGVAVGTPTALLVAGVTGLISGLEASKQAMFESVANRLOGKILEWEK 407
Db 368 LAATAGGVSAAGVAVASPTALLVSGITGVISTILQYSKQAMFHVANKHKNKIVWEK 427
Qy 408 QNGGQNYFDKGYDSRYAAYLANNKFLSELNKELEAERVIAITQOORNNIGELAGITKL 467
Db 428 NNGKNVFEYGDYARYLANLQDNKFLNLNKELOAERVIAITQOOWDSNIGDLAGISRL 487
Qy 468 GERIKSGRAYADAFEDGKVKVAGSNITLDAKTGIIIDISNSNGKKTQALHFTSPLLTAGE 527
Db 488 GEKVLGKAYVDAFEQGHKADKLVLQDSAKGIIIDVSNTEAKTQHILFRTPELLTPTGE 547
Qy 528 SRERLTNGKYSYINKLKFGRVKNQVTDGEASSKLDLFSKVIQRA-----ETEGTDEI 580
Db 548 KRERVQTKYBYITKLHINRVDSWQIKDGAASSTFDLTNNVQRIQVGLDHAENVIKET 607
Qy 591 GLIYNVAKAGNDDIFVQCGKMNIDGGDGHDRVYFSKGGFGNITVDGTSATAGSYTVNRK 640
Db 608 KIVATLGGDDNVFVGSQTTIDGEGYDRVHYSR-GNYGALTIDATKETQGSYTVNRP 666
Qy 641 VARGDIYHEVVKRQETKVKRTETITQYRDYELRVKGYGYQSTDLNLSVEEVIGSQFNDVF 700
Db 667 VESGKALHEVTSHTALVGNREEKIEYR-HSNQHHAGYYTKDTLKAVEEIIIGTSHNDF 725
Qy 701 KGSFENDIFHSGEGDLDLGGAGDRLFGKGNDRLSDEGDDLLDGGSGDDVNLGGAGN 760
Db 726 KGSFENDAFNGDGVDTTIDGNDGNDRLFGKGGDDIIDGGNGDDFTDGGKGNLLHGGKGD 785
Qy 761 DVYIFRKGDNLTLDYDGTGNKLFADANISDIEMERTKEGIIKVRNDHSGSINIPRWY- 819
Db 786 DIFVHRQGDGNDSTESGNDKLSFSDSNLKDLPFEKYNHILVI-TNTQEKVYIQNWR 844
Qy 820 ---ITSNLQYQSNKTHKIEQLIGKQSYITSQIDKIQKDGTVITSOELKKLADE 876
Db 845 EAEPAKTIRNVATR-DDKIEIIGQGERITSQVDELIBKKG--KIDKSDLSSQVVDN 901
Qy 877 NKSQKLSASDTSASLNKLGVSNALFGTANSVSSNALQITOPTQGIAPSV 927
Db 902 YQLLKYS-RDASNSLDKLSSASAFSTSSND-SRNVL---ASPT-SMLDPSL 946

RESULT 19
Q93NPI PRELIMINARY; PRT; 956 AA.
AC Q93NPI;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
RTX toxin IIA.
DE Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=715;
RN [1]
RP SEQUENCE FROM N.A.
RA Shin S.-J., Park J.-Y., Choi I.-S., Shin N.-R., Yoo H.-S.;
RT "Cloning and Sequencing of Apx IIA from Actinobacillus
RT pleuropneumoniae."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; AF363362; AAK50052.1; -.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 5.
DR Pfam; PF02382; Rtx; 1.
DR PRINTS; PRO0313; CABNDNGRPT.
DR PRINTS; PRO1488; RTXOKINA.
DR PROSITE; PS00330; HEMOLYSIN CALCIUM; 1.
DR PROSITE; PS00330; HEMOLYSIN CALCIUM; 1.
SQ SEQUENCE 956 AA; 102495 MW; 9F8EB14EC76628D5 CRC64;

Query Match 48.0%; Score 2229.5; DB 2; Length 956;
Best Local Similarity 47.7%; Pred. No. 1.1e-91;
Matches 453; Conservative 187; Mismatches 25; Indels 53; Gaps 15;

Qy 1 MSNINV--IKSNIOAGI-----NSTKSGLNKLALAIKPDYDQ 36
Db 1 MSKITSSLSKSSIQQGLKQKQKLNQAGTTLKNGLTQTHSLQNGAKLLIYIPOGYDSG 60
Qy 37 KGGLTNDFIKAADDELGIARLAEBPNHTETAKKSVDTVNOFLSLTQTGTGIAISATKLEKPLQ 96
Db 61 QGNGVQDLVKAANDLGIENVREERSNLDIAKTSFDTTQKILGFTDRGIVLFPQLDNLK 120
Qy 97 KHSTNKLAGLDSVENIDRKLGKASNVLSLSPFLGTALAGIELDSLIKKGDAAPDALAK 156
Db 121 KNP--KIGNTLGSASSISQNIQKANTVILGGIQLSGVLSVNLNELLQNKDPNQLLEAK 178
Qy 157 ASIDLNEIIGNLSQSTQTTIEAFSSQIAKLGSTISQAKGFSNIGKLNQNL-NFSKTNLGL 215
Db 179 AGLELTNELVGNLASSVQTVDAFAEQISKLGSHLQNVKGLGGLSNKLNQNLPLDKASLGL 238
Qy 216 EITLGLSGISAGFALADKNASTGKKVAAGFELSNOVIGNVTKAISYVLAQRAAGLST 275
Db 239 DIISGLLSGASAGLLADKEASTEKKAAAGVEFANQIIGNTVKAIVSSVILLAQRAVAGLSS 298
Qy 276 TGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQGVG 335
Db 299 TGPVAALIASTVALAVSPLSFLNAVADFQKQADLIKSYSERFQKLGVDGDRLLADFHRETG 358
Qy 336 TIEASLTITFTALGAVSAGVSAAGVAVGTPTALLVAGVTGLISGLEASKQAMFESVA 395
Db 359 TIDASVTITNTALAAISGGVGAASAGSLVAPVALLVAGVTGLITLITILEYSKQAMFEHVA 418
Qy 396 NRLOGKILEWEKQNGGQNYFDKGYDSRYAAYLANNKFLSELNKELEAERVIAITQQRWD 455
Db 419 NKVHDRVIVEKKH-NKNYFEQGYDSRHLADLQDNMKFLNLNKELOAERVIAITQQRWD 477
Qy 456 NNIGELAGITKLGERIKSGAYADAFEDGKVKVAGSNITLDAKTGIIIDISNSNGKKTAL 515
Db 478 NQIGDLAAISRRTDKISGKAYVDFAFEQGHQSYDSSVQDLNKNKGIINISNTN-RKTQSV 536
Qy 516 HFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNQVTDGEASSKLDLFSKVIQRA--- 572
Db 537 LFRPLITPGEENRERQEGKNSYITKLHQRVDSVITVDGASSVDFTNVVQRIAVKF 596
Qy 573 -----ETEGTDEIGLIVNAKAGNDDIFVQCGKMNIDGGDGHDRVYFSKGGFGNITVD 625
Db 597 DDAGNIIESKDTK--IIANLGAQNDNVFVGSSTTVIDGGDGHDRVHYSR-GEYCALVID 652
Qy 626 GTSATEAGSYTVNRKVARGDIIYHEVVKRQETKVKRTETITQYRDYELRVKGYQSTNL 685
Db 653 ATAETEKGSYVKRYVGDSCALHETIATHTPTNVGNREEKIEYR-REDDRFHTGYTVTDSL 711
Qy 686 KSVEEVIGSQFNDVFKSGKENDIFHSGEGDLDLGGAGDRLFGKGNDRLSDEGDDLL 745
Db 712 KSVSEIIGSQNDNIFKGSQFDDVFGNGVDTIIGNDGDHDLFGAGDDVDIDGGNGNLF 771
Qy 746 DGGSGDDVNLGGAGNDVYIFRKGDNLTLDYDGTGNKLFADANISDIEMIERTKEGIIIVK 805
Db 772 VGGTGNDIISGKNDIYVHKTDGNDSDITDSGGQDKLAFSDVNLKDLTFFKVDSSLEI- 830
Qy 806 RNDHSGSINIPRWY-----ITSNLQYQSNKTHKIEQLIGKQSYITSQIDKIQKDG 861
Db 831 INQKGEKVRIGNWFLKNDLASTVANYKAT-NDRKIEEIIIGKGBRITSQVDKLI--KEG 887
```

```
Qy 862 GTVITSQELKADENKSKLSASDIASSLNKLVGSMALPGTANSVSSN 910
Db 888 NNQISAKALSQVGNQYNTSK-DRONVSNLSAKLISSVESFTSSNFRNN 935

RESULT 20
Q8KW26
ID Q8KW26 PRELIMINARY; PRT; 987 AA.
AC Q8KW26;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE AQXA.
GN AQXA.
OS Actinobacillus equuli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=90-570;
RX MEDLINE=22031201; PubMed=12034544;
RA Berthoud H., Frey J., Kuhnert P.;
RT "Characterization of Agx and its operon: the hemolytic RTX determinant
of Actinobacillus equuli.";
RT Vet. Microbiol. 87:159-174(2002).
DR EMBL; AF381185; AAM45569.1; -.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 3.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 1..
SQ SEQUENCE 987 AA; 105871 MW; 35440030AB522B53 CRC64;

Query Match 46.7%; Score 2171; DB 2; Length 987;
Best Local Similarity 48.1%; Pred. No. 5e-89;
Matches 468; Conservative 163; Mismatches 264; Indels 78; Gaps 23;

Qy 3 NINV-----IKSNIAQGLNSTKSGI-----KN-----LYLAIPROYDPQKGGTIND 43
Db 17 NLNLQFOQGAGKISGVQKQVQSSQKLSQAGATKNVANKILYIPKEYDATAGNLKG 76

Qy 44 FIKADELGTARLAEBENHTETAKKSVDTVNOFLSTQTGTIAISATKLEKLOKHSNTKL 103
Db 77 LLDAEALGIEVQREKDGATAYSAFVGVEQFTGLTERGTLTLPQLDNFLKHP--KF 134

Qy 104 AKGL-DSVENTDRKLGKASNVLSLTSSPLGTALAGIELDSLKKGDADAPALAKASIDL 162
Db 135 SKSGDSSEKVLNVGNKILSGISQVLTGLTLAGIELDNLKEGGSKTE-LAQAGINLV 193

Qy 163 NEIIGNLSQSTQTTIEAFSSQLAKLGSTISQAKFSNIGNKLNINL---FSKTNLGLBIIT 219
Db 194 NELVGNIAKGAETIQAFSDQLAHLGVSIVQNTKGLGGLGNKLNISGSAKAGLGFDIVS 253

Qy 220 GLLSGISAGPALADKNASTGKVAAGFELSNOVIGNVTKALSSVTLQORVAAGLSTTGAV 279
Db 254 SLLSGVAGFALADPNASTGKIAAGFELSNOVIGGITKAVSGVILAQRIASGLSTTGPA 313

Qy 280 AALITSSIMLAISPLAFNNAADKFNHANALDEPAKQFRKFGYDGDHLLAEVQGVGTIEA 339
Db 314 IGLTASSISLAISPLAFNLVADRPHANEINKFAERFKFGYEGDSLLADFHESGAIDA 373

Qy 340 SLTTISTALGAVSAGVSAAGVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQ 399
Db 374 SITTTINTALGAISAGVTAAGASLVGAPVTLVGGITGLISAILDLSKQAMFEHVATKLS 433

Qy 400 GKILEWKQNGQNYFQKGYDSRVAAYLANNLKFLSELNKELEARVIAITQQRWDNNIG 459
Db 434 GKIEAWKKY-GMNYFEKGYDARHAAFLIEDNFAFESLITKELRAERVISITQQQWDAQIG 492

Qy 460 ELAGITKLGRIKSGKAYADAFEDGKKVEAGSNI-TLDAKTGIIDISNSNGKKTQALHFT 518

Db 493 DLAGITRGDKIQSGKAYVDVFPKESKEDRNDYNIIVDFPTEGIIIDISKTT-SKTQHLLFL 551
Qy 519 SPLLTAGTSERRLTNGKYSYINKLAFGRVKNQVTDGEASSKLDPSKVIQORVAETGTD 578
Db 552 NPLLTPEKENRERREKKGYEYVTKLIIVDRKTQVQVTDGEASSSTLDFTNVLQFIA--VDTD 609
Qy 579 EIGLI-----VNAK--AGNDDIFVGOGRMNIDGGDGHDRVFPYSKDGGFNGI--TVDG 626
Db 610 RAGNVTESLEAKTEAKLKGDDTVFVGWSTDIDGSEGYDRAAYNRDLDEGHIYGLNIDA 669
Qy 627 TSATEAGSYTVNRKVARGDIIYHEVVKRQETKVGKRTTETIOYRDYELR-----KVGYGYQS 681
Db 670 QQETVAGSYTVNRTICGGAKEVEIKVHQATACKRVDKIEYRQGESRFHMLKV-----723
Qy 682 TDNLKSVVEVIGSQFNDVPKSGKFNDI FHSGECDLLDGGAGDRLFGGKGNDRLSGDEG 741
Db 724 VDKLVNVEEIIIGTKNQDI FPGSKFNDTFFGGDEKDIYGNAGNDRLFGGNGNDEIDGGDG 783
Qy 742 DDLDDGGSDVVLNGAGAGNDVYIFRKGDDGNDTLYDGTGNDKLAFAFADANISDIERTKEG 801
Db 784 NDFIDGKGNDSLTCGYGNDIYVIHKEDGQDFITELGGDRLVFTDTLSRLTFKRDKV 843
Qy 802 IIVKRNHSGSINIPRWYI-----TSNLQNYQSN--KTDHKIEQLIGKSGSYITSQIDIKI 855
Db 844 LIITEDSAHNNQVRISDWFFKQATLNKEYHAQSAKDVKYRIEIIDKNGTRITADQIDTI 903
Qy 856 LQDKKDGTVITSQELKADENKSKLSA-----SDIASLNKLVGSMALP-----G 902
Db 904 LNKSGEIVESQ-LVKISDDYKA-KVAADNAHLSEIAEALGKIASASSASFSAVTGGAMG 961
Qy 903 TAN---SVSSNAL 912
Db 962 SANAFPAVSENAL 974

RESULT 21
Q8KW29
ID Q8KW29 PRELIMINARY; PRT; 987 AA.
AC Q8KW29;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE AQXA.
GN AQXA.
OS Actinobacillus cf. equuli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=182091;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCUG 19799;
RX MEDLINE=22031201; PubMed=12034544;
RA Berthoud H., Frey J., Kuhnert P.;
RT "Characterization of Agx and its operon: the hemolytic RTX determinant
of Actinobacillus equuli.";
RT Vet. Microbiol. 87:159-174(2002).
DR EMBL; AF381184; AAM45566.1; -.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 1.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 1..
SQ SEQUENCE 987 AA; 105662 MW; CC34CD62BF9B745D CRC64;

Query Match 46.4%; Score 2155; DB 2; Length 987;
Best Local Similarity 47.5%; Pred. No. 2.6e-88;
Matches 463; Conservative 167; Mismatches 266; Indels 78; Gaps 23;

Qy 2 SNINV-----IKSNIAQGLNSTKSGI-----KN-----LYLAIPKDYDPQKGGTIN 42
Db 2 SNINV-----IKSNIAQGLNSTKSGI-----KN-----LYLAIPKDYDPQKGGTIN 42
```

Db 16 SNLNLQFOGAGKISGVQKQSQHLSQAGAAATKNYANKLJLYIPKEYDATAGNGLK 75
Qy 43 DFIKAADDELGTARLAEBEPNHTETAKSVDTVNOFLSLTQTGIAISATKLEFLOKHSNK 102
Db 76 GLLDAAEALGIEVOREEDGVTATAYSAFVGEOQIFGLTERGLTFLPQLDNFLKHP--K 133
Qy 103 LAKGI--DSVENIDRLKGRASNVLTSLSSFLGTALAGIELDSLIIKKGDAAPDALAKASIDL 161
Db 134 FSKSIGDSEKVLGNVGNNAKILSGIQSVLGTTLAGIELDNLIEKGGSKTB--LAQAGINL 192
Qy 162 INEIGNLSQSTQTIEAFSSQAKLIGTSIOAKGFSNIGNKLQNLN---FSKTNLGLIEII 218
Db 193 VNELVGNIAKGAETIQAFSDQLAHLGSVIQNTKGLGGLGNKLQNLISGSALNAKAGLGFIV 252
Qy 219 TGLISGISAGPALADKNATKGVAAAGPELSNOVIGNVTKAISYVLAQRAAGLSTTGA 278
Db 253 SSSLISGSAGPALADPNASTGKTKAAGPELSNOVIGGITKAVSGVILAQRIASGLSTTGP 312
Qy 279 VAALITSSIMLAISPLAFWNAADKENHANALDEFAKQFRKGYDGDHLLAEYQGVGTIE 338
Db 313 AIGLIASSISLAISPLAFWNAADKENHANALDEFAKQFRKGYDGDHLLAEYQGVGTIE 372
Qy 339 ASLTTISTALGAVSAGVAAAAGVAGTPIALLVAGVTGLISGILEASKQAMFESVANRL 398
Db 373 AAITTINTALGAISAGVTAAGASLVGAPVILLVGGVTGLISAILDLSQAMFEHATKL 432
Qy 399 OGKILEWKONGQNYFQKGDSDRSVAAYLANNLKPLSELNKELEBAERVAITQORWNNI 458
Db 433 SGKIEWBEKKY--GMNFEKGYDARHAAFLDNFAFFESLTKELAERVISITQOQWDTQI 491
Qy 459 GELAGITKGERIKSGKAYADAFDQKKVEAGSNI--TLDAKTGIIIDISNSNGKKTQALHF 517
Db 492 GDLAGITRRGKIQSGKAYVDVFKETKDLSDYDNIIVTDPTEGIIIDISKT--SKTQHLLF 550
Qy 518 TSPLLTAGTRESRLTNGKYSINKLKFRGVKNQVTDGEASSKLDLFSKIQIRVAETEGT 577
Db 551 LNPLLTGPKENPEREKKGYEYVTKLIIVDRKTKWQVTDGEASSTLDTFNVLIQFIA--VDT 608
Qy 578 DEIGLI-----VNAK--AGNDIDIFVGGKKNIDGGDGHDRVFKSGGFGNI---TVD 625
Db 609 DRAGNVTESLEAKIEAKLKGKDDTVFVGWGSTDIDGEGYDRAAYNRILDGEAHYGLNID 668
Qy 626 GTSATEAGSYTVNRKVARGDIVHEVVKQETKVGKRTETIQYRDVELR-----KVGYGQ 680
Db 669 AQQETVAGSYTVNRITAGAAKHVIVKHQATAGKRVKIEYRQGEQRYHODLVK----- 723
Qy 681 STDNLKSVEEVIGSFQNFVFKSGKENDIFHSGEGDLDLGGAGDRLFGGKGNDRLSGDE 740
Db 724 -VDKLLNTEIEIGTKNRDNFKSGKENDIFFGGDEKDIIDGNAGNDRLFGGDEIDGDD 782
Qy 741 GDDLGGSGDDVLNGGAGNDVYIPRKGNDTLVDGTDGNKLAFAADANISDIWIERTKE 800
Db 783 GNDFIDSGKDDIISGCHGNDIYVHKDGAQDIYTELGGDDRLVTDTELSRLTFKRDNR 842
Qy 801 GIIVKRNHSGSINTPRVYISNL---QNYQSNKT---DHKIEQLIKGGSVITSQDIDK 854
Db 843 TLVIEDSANNHVQVSDMFLKQDQAKINKEYHAQSVLDKDYRTEIIDKNGTRITADQIDK 902
Qy 855 ILQDKKDGTVITSQELKKLADENKQKLSA-----SDIASSINKLVGSMALF----- 901
Db 903 ILNGSEGVIEVSQ--LVKISDDYKA--KVAADNAHLSEIAEALGIASSASFSAVTGAM 960
Qy 902 GTAN---SVSSNAL 912
Db 961 GSANAFAAVSENAL 974

RESULT 22
Q9RCG8
ID Q9RCG8 PRELIMINARY; PRT; 1049 AA.
AC Q9RCG8;
DT 01-MAY-2000 (TreMBLrel. 13, Created)
DT 01-MAY-2000 (TreMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TreMBLrel. 23, Last annotation update)
DE PaxA.
GN PaxA.
OS Pasteurella aerogenes.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=749;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=field strain JFI319;
RA Heyberger-Meyer B., Frey J., Nicolet J., Kuhnert P.;
RT "Identification and characterization of a new RTX toxin operon (pax)
in Pasteurella aerogenes";
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U65588; AAF15370.1; --
DR InterPro; IPR001343; Hemolysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF02353; hemolysinCabin; 6.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PS00330; HEMOLYSIN_CALCIIUM; 1.
DR PS00330; HEMOLYSIN_CALCIIUM; 1.
SQ SEQUENCE 1049 AA; 112308 MW; 9FA5070B48CC3127 CRC64;

Query Match 43.1%; Score 2004.5; DB 2; Length 1049;
Best Local Similarity 43.7%; Pred. No. 1.5e-81;
Matches 439; Conservative 158; Mismatches 288; Indels 119; Gaps 20;

Qy 6 VIKSNIQAGLNSTKSGLK-----NLYLAIPKDYDPQKGGTIND 43
Db 17 IAKQAKKGIDVTNKGLQYGVSVQKLQALAAKRSIQYGNKLVLPKDYDVTNGGFFD 76
Qy 44 FIKAADELGLTARLAEBEPNHTETAKSVDTVNOFLSLTQTGIAISATKLEFLOKHSNK 103
Db 77 LAKAAEELGIVQKYIDRNDLEIAHKSGLVTDQFGLTERGLTFLPQLDKFLOQHSKISN 136
Qy 104 AKGLDSVENIDRLKGRASNVLTSLSSFLGTALAGIELDSLIIKKGDAAPDALAKASIDLIN 163
Db 137 VGSSTGDTVN--KLAKSQAIIISGVSVLGINLNEALISGSELE--LAKAGVDLAS 194
Qy 164 EIIGNLSQSTQTIEAFSSQAKLIGTSIOAKGFSNIGNKLQNLN---FSKTNLGLIEITG 220
Db 195 ELVGNIAKGTATIEAFSEIQNFQKLVQNAKGLGGVQQLQHIISGALSSTGLGLDIIS 254
Qy 221 LLSGISAGPALADKNATKGVAAAGPELSNOVIGNVTKAISYVLAQRAAGLSTTGA 280
Db 255 LLSGVTASTLADKNASTSTKVAAGPELSNOVIGGITKAVSSYTLAQLAAGLSTTGPAA 314
Qy 281 ALITSSIMLAISPLAFWNAADKENHANALDEFAKQFRKGYDGDHLLAEYQGVGTIEAS 340
Db 315 ALIASSISLAISPLAFWNAADKENHANALDEFAKQFRKGYDGDHLLAEYQGVGTIEAS 374
Qy 341 LTTISTALGAVSAGVAAAAGVAGTPIALLVAGVTGLISGILEASKQAMFESVANRLQ 400
Db 375 ITTISTALSAIAAGTAAASAGALVGPITLLVGTIGLISGILEFSKQPMLEHVASLGT 434
Qy 401 KILWEKONGQNYFQKGDSDRSVAAYLANNLKPLSELNKELEBAERVAITQORWNNICE 460
Db 435 KIEWERKY--GKNYFENGIDARHKAFLSDLSLSSFNKQYETERAVLITQORWDEYIGE 493
Qy 461 LAGITKGERIKSGKAYADAFEDG---KKVEAGSNTLDAKTGIIIDISNSNGKKTQALH 516
Db 494 LAGVTGKDKLISGKAYVDIPEEGKLLAKRPDDFNRLVLDPKKIDISNS--QTSTLLK 551
Qy 517 FTSPLLTAGTRESRLTNGKYSINKLKFRGVKNQVTD--GEASSKLDLFSKVIQR----- 570
Db 552 FVTELLTPGTESRKRRTQTKGYEYVTKLVNVDNGINQWVNGVKEKAVYDFTNLIQHVHIS 611
Qy 571 -VAETEGDEIGLIVNAKAGNDIDIFVGGKKNIDGGDGHDRVFKSGGFGNITVDGTS 629
Db 612 SVARGEYREVLVSRILSGKNDKVFLLASGSABIHAGDGHVYVYDKT--DTGLLWVDGQA 670
Qy 630 TEAGSYTVNRKVARG--DIYHEVVKRQETKVGKRTETIQYRDVELRKVG--YGVQSTDNLKS 687

Db 671 TKQGYTVTRELSGATQILREVVRKQKSSVGRQETVEYRDELNAQSGSNLAKADNLYS 730
Qy 688 VEEVIGSFQNDVFKGSKFNDIFHSGEGDLDL-----G 720
Db 731 VEEIIGSNHRDEFKGSKFRDIFHGADGGDLDLNGNDGDDILYCDKGNDELNGNDQLYG 790
Qy 721 GAGDRLFGGKGNDRLSGDEGD-----DLDDGGSGDD 752
Db 791 GEGNDKLFPGNGNYLSGGDGDDELQVNGFNVLRGKGNDKLYGGAGSDFLDGGEGDD 850
Qy 753 VLNGAGNDVYIFRKGNDNTLYD---GTGNDKLPADANISDIEMIERTKEGIIIVKRNH 809
Db 851 YLAGGEGNDVYVYRSTGSHHTIYDQKSSDSTLYLSDLTDRLLAVEKVDNVLVKPSDH 910
Qy 810 S---GSINI PRWYITSNLQYQSKNDIHIKQLGKDGSYITSQIDKILQKDKGTVIT 866
Db 911 NSNRGSLTIKDWFTKGHYN-----HKLQIIVDKNGRKLTSN-----LETHFNTPKT 959
Qy 867 SQELKCLADENKSKLASDIASSLNKLVGSMALPGTANSVSN 910
Db 960 N--LLGYTAQNQNES--NLSSKLTGKIISAGNPLAKQGNNN 1000

RESULT 23

Q93NPO PRELIMINARY; PRT; 1052 AA.
AC Q93NPO;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE RTX-toxin IIIA.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=715;
RN [1]
RP SEQUENCE FROM N.A.
RA Shin S.J., Park J.Y., Choi I.S., Shin N.R., Yoo H.S.;
RT "Cloning and Sequencing of Apx IIIA from Actinobacillus
RT pleuropneumoniae";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF363363; AAK50053.1; -
DR InterPro; IPR001343; Hemlyen_Ca_bind.
DR InterPro; IPR003995; RTX.
DR Pfam; PF00353; HemolysinCbind; 6.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CAENDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN CALCIUM; 1.
SQ SEQUENCE 1052 AA; 112862 MW; CA0E160F02101F42 CRC64;

Query Match 42.1%; Score 1956.5; DB 2; Length 1052;

Best Local Similarity 42.6%; Pred. No. 2.1e-79;

Matches 425; Conservative 167; Mismatches 286; Indels 119; Gaps 19;

Qy 8 KSNTOAGLNTSKSLK-----NLYLATPKDYDPQKGTLDNFI 45
Db 19 KRQAKGYDVTNKLQGVSVQAKLQALAGKAVQYGNKLVLPKGYDSVGVNGFFDLV 78
Qy 46 KAADDELGTARLAEEPNNHTAKKSVDTVNFSLTGTGIAISATKLEKFLQKHSTNKLAK 105
Db 79 KAARELGQVYVNRNEVAHKSIGTADQFLGTERGLTSLAPQLQFLQKHKSINVV 138
Qy 106 GLDSVENIDRKLKASNVSLTSSFLGTLGAGIBLDSLKKGDAAAPALAKASIDLINVI 165
Db 139 GSSTGDVAVS-KLAKSQTIIISQISVGLTGLAGINLNEALISGSGSELE-LAEAGVSLASEL 196
Qy 166 IGNLSQSQTOTTEARSSQLAKLGSITSQAKGSNTGNKLNLN---FSKTNLGLBIIITGLL 222
Db 197 VSNIAKGTITTDAPTQIQNFQKLVENAKGGVGRQLQNTSGSALSKTGLGLDIISLL 256
Qy 223 SGISAGPALADKNASTGKVAAGFELSNOVIGNVTNKAISSVYLAQRVAAGLSTTGAVAL 282

Db 257 SGVTASALANKNASTTKVAAGFELSNOVIGGITKAVSSVYLAQRVAAGLSTTGAVAL 316
Qy 283 ITSSIMLAISPLAPMNAADKFNHANALDEFAPKQRFKFGDGDHLLAEYQVGVGTIEASLT 342
Db 317 IASSISLAISPLAPLRVADNFRNSKETGEFAERFKLGYDGDGKLLSEFYEAGTIDASIT 376
Qy 343 TISTALCAVAGSAGVSAAGVAGTPIALLVAGVTGLISGLEASQKQAFESVANRLOSKI 402
Db 377 TISTALSAIAGTAAGTAAGALVGPITLLVTGITGLISGLEFSPKQPMPLDHVASKGNKI 436
Qy 403 LEWKKQNGQNYFDKGYDSRYAAYLANNLFLSELNKELEAERVIAITQQRWDNNIGELA 462
Db 437 DEWEKKY-GKNYPENGVDARHKAFLSDSFLSLSNFKQYETERAVLITQQRWDEVIGELA 495
Qy 463 GITKLGRIKSGKAYADAFEDGKKVAG-----SNITLDAKTGIIDISNSNGKKTQALHPT 518
Db 496 GITGKDKLSSGKAYVDYFQEGKLEKPPDFSKVDFPTKGEIDISNS--QTSTLLKFV 553
Qy 519 SPLLTACTESRERLTNCKYSYINKLKFRVKNQVOTDG--EASSKLPDFSVIOR----- 570
Db 554 TPLTPTGESRERTQTKYETITKLVVKGDKM--VNVGVKDKGAVDYTNLIQHAHISSS 612
Qy 571 VAETEGTDEIGLIVNAKAGNDIDFVGQGNNDIGDGDGHDHVRVYKDGCGFGNITVDGTSAT 630
Db 613 VARGEYREVRLVSHLGNLGNKDVFLAAGSAEIHAGEGHVVYDKT-DTGLLVIDGTAT 671
Qy 631 EAGSYTVNRKARG-DIYHEVVRQETKQKRTETIYROYELRKVG-YGVQSTDNLKSV 688
Db 672 EQGRYSVTRELSGATKTLREVINKQKSAVGKREETLEYRDTYELTQSGNSNLKADELHVS 731
Qy 689 BEVIGSQFNDVFKGSKFNDIFHSGEGDLDLGGAGDRLFGKGNDRLSGDEGD----- 742
Db 732 EEIIGSNQRBEPFKGSKFRDIFHGADGDLNGNDGDDILYDCKGNDELNGNDQLYGG 791
Qy 743 -----DLDDGGSGDDV 753
Db 792 EGNDKLLGGNGNYLSGGDGNDELQVLGNGFNVLRACKGDDKLYSGSGSLLDGGEGNDY 851
Qy 754 LGGAGNDVYIFRKGNDNTLYD---GTGNDKLPADANISDIEMIERTKEGIIIVKRNH 810
Db 852 LEGDGSDFYVYRSTGSHHTIYDQKSSDLDKLYLSDFSFRLAVEKVDNVLVRSNESS 911
Qy 811 GS---INIPRWYITSNLQYQSKNDIHIKQLGKDGSYITSQIDKILQKDKGTVITS 867
Db 912 HNNRVLTIKDWFT-----KEGNKNYKHKEIQIVDKNGRKLTAENIGTYFKNAPKADNLLN 964
Qy 868 QELKCLADENKSKLASDIASSLNKLVGSMALPGTA 904
Db 965 YATKE--DQNES---NLSSKLTLSKLTITNAGNFGVA 996

RESULT 24
Q8FE01 PRELIMINARY; PRT; 1024 AA.
AC Q8FE01;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Hemolysin A.
GN HLYA OR C3570.
OS Escherichia coli O6
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22398234; PubMed=12471157;
RA Raeko R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;

RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic *Escherichia coli*.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016766; AAN82018.1; -.
KW Complete proteome.
SQ SEQUENCE 1024 AA; 110246 MW; 9DEBA50C23B5A49C CRC64;

Query Match 39.8%; Score 1848.5; DB 16; Length 1024;
Best Local Similarity 41.8%; Pred. No. 1.4e-74;
Matches 428; Conservative 153; Mismatches 310; Indels 132; Gaps 25;

QY 2 SNINVIKSIQAGLNSTKSLK-----NLYLAIPKDYDPQKGGTTLDFKAADE 50
DB 19 SSANKLHS---AG-OSTKDALKAABQRTNAGNRLILLIPKDYKQ-GSSLNDLVRTADE 73
QY 51 LGIARLAEPNHTETAKSVDTVNOFLSLTQTGTAISATKLEKFLQKH--STNKLAKGLD 108
DB 74 LGIEVQYDEKNGTAITKQVFGTAELKLGITRGVTIFAPQLDKLLQKYKAGNKLGG--G 130
QY 109 SVENIDRLKASNVLSLSSFLGTALAGIELDSLKK-----GDAAPDALAKASIDLINE 164
DB 131 SAENIGDNLGKAGSVLSTFQNLGALSSMKIDELIKRKQSGSNVSSSELAKASIELINQ 190
QY 165 IIGNLSQSTQTEAPFSOLAKGSTISQAKGFSNIGNKLQNL-NFSKTNLGLIEITGLLS 223
DB 191 LVDTPAASINNUNVNSFSQNLKGLSVLNTKLTGVGNKQLNLPNLNDIGAGLDTVSGILS 250
QY 224 GISAGFALADKNASTGKVAAGFELSNOVIGNVTKAISYVLAORVAAGLSTTGAAALI 283
DB 251 AISASFILSNADADTGCTAAAGVELTTKVLGNVKGISQYIIAQAQGLSTSAAGLI 310
QY 284 TSSIMLAISPLAFMAADKFNHNAALDFAPQKFRFGYDGHLLAEYORGVTIEASLTT 343
DB 311 ASVVTLAISPLSFLSIADKFRANKIEYSQRFKLGVDGSLAAAFHKTGAIDASLTT 370
QY 344 ISTALGAYSAGVSAAGVSGVPTIALLVAGVTGLISGILEASKOAFESVANRLQKIL 403
DB 371 ISTVLASVSSGISAATTSLVGAPVSLVAGVTGLISGILEASKOAFESVANRWADVIA 430
QY 404 EWEKONGQNFDPKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQORWNNNIGELAG 463
DB 431 EWEKXH-KKNYFENGYDARHAFLADNFNFKLSQYNKEYSVSVSLVITQOHWDTLIGELAG 489
QY 464 ITKLGERTKSKAYADAFEDGKKVEAG-----SNITLDAKTGIIDISNSGKKTQALHFTS 519
DB 490 VTRNGDKTLGSKSYIDYEEGKRLEKPKDFEQKQVDFPLKGNIDLSDS--KSSTLLKEVT 547
QY 520 PLTAGTESRRLTNGKYSYINKLKFRGVKNQVTD-GEASSKLDKSVIQRVAETG- 577
DB 548 PLITPGEIRERROSGKYEYITELLVGVKWKVGVODKGSVYDYNLIQHASVGNQY 607
QY 578 DEIGLIVNAKAGNDIDFVQGGKMNIDGGDHRVFSK-DGFGNITVDGTSATEAGSYT 636
DB 608 REIRTESHLGDGDDKVLFSAGSANIYAGKGDHVVYDKTDG--LTIDGKATEAGNYT 665
QY 637 VNRKVARGD--IYHVEVKROETKCKRTETTOYRDYELRKV-GYCYQSTDNLKSVEEVIG 693
DB 666 VTR-VLGGDVTKLQEVWEQEVSQKRTKQYRSYEFTHNGKNLTETDNLYSVEELIG 724
QY 694 SQFNDVFKGSKFNDIFHSGEGDDLDGCGADDRLF----- 728
DB 725 TTRADKFPGSKFTDIFPHGADGDHDEGNDGNDRLVGDKNDTLGGNGDDQLYGGDGNK 784
QY 729 -----GKGNDRLSGDEGDDLLDGGSGDDVNLG 756
DB 785 LIGTGNNYNGCGDDELQVQGNLAKNVLGSGKGNKLYGSEGADLLDGGEGDNLKG 844
QY 757 GAGNDVYIFRKCDGNDTLYDGTG-NDKLAFAADANISDMIERTKEGIIVKRND----- 808
DB 845 GYGNDIYRLSGYGHIIIDDDGGKDKLSLADIDFRDVAFRREGNDLIMYKAEGNVLSIG 904
QY 809 HSGSINIPIRWYITSNLQYNSNKTDKIEQLIKGDSYITSDIKILQDKK----- 860

DB 905 HKNGITFRNWF-----EKESGDISNHQIEQIDKXGRVITPDSLKLALEYQOSNNKASYV 959
QY 861 ---DGVITTSQ-----ELKKLADENKSKQLSASDIASLNKLVSMALEFCTA-NSVS 908
DB 960 YGNDALAYSQDNLNLINEISKIISAAGNFDVKEERAASALLQLSGNDSFSGYRNSIT 1019
QY 909 SNA 911
DB 1020 LTA 1022

RESULT 25
Q8GA40 PRELIMINARY; PRT; 1024 AA.
AC Q8GA40;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hemolysin A.
GN HLYA.
OS *Escherichia coli*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; *Escherichia*.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=536;
RX MEDLINE=22267134; PubMed=12379716;
RA Dobrindt U., Blum-Oehler G., Nagy G., Schneider G., Johann A.,
RA Gottschalk G., Hacker J.;
RT "Genetic structure and distribution of four pathogenicity islands (PAI
1536 to PAI IV536) of uropathogenic *Escherichia coli* strain 536.";
RL Infect. Immun. 70:6365-6372 (2002).
DR EMBL; AJ488511; CAD33759.1; -.
SQ SEQUENCE 1024 AA; 110314 MW; 9771FOA61ECF6EEB CRC64;

Query Match 39.7%; Score 1845.5; DB 2; Length 1024;
Best Local Similarity 42.1%; Pred. No. 1.9e-74;
Matches 424; Conservative 153; Mismatches 303; Indels 128; Gaps 25;

QY 17 STKSLGK-----NLYLAIPKDYDPQKGGTTLDFKAADELGIAEENHFTET 65
DB 30 STKDALKAABQRTNAGNRLILLIPKDYKQ-GSSLNDLVRTADELGIEVQYDEKNGTAI 88
QY 66 AKKSVDTVNOFLSLTQTGTAISATKLEKFLQKH--STNKLAKGLDSVENIDRLKASNV 123
DB 89 TKQVFGTAELKLGITRGVTIFAPQLDKLLQKYKAGNKLGG--GSAENIGDNLGKAGSV 145
QY 124 LSTLSSFLGTALAGIELDSLKK-----GDAAPDALAKASIDLINEIIGLSQSTQIEAF 179
DB 146 LSTQNFGLTALSSMKIDELIKRKQSGSNVSSSELAKASIELINQVDTAASINNUNVNSF 205
QY 180 SSQALAKGSTISQAKGFSNIGNKLQNL-NFSKTNLGLIEITGLLSGISAGFALADKNAST 238
DB 206 SSQLNKLGSVLSNTKHLNGVGNKQLNLPNLNDIGAGLDTVSGILSVISASFILSNADAT 265
QY 239 GKVAAGFELSNOVIGNVTKAISYVLAORVAAGLSTTGAAALITSSIMLAISPLAFMN 298
DB 266 GTKAAAGVELTTPKVLGNVKGISQYIIAQAQGLSTSAAGLITASVTLAISPLSLS 325
QY 299 AADKFNHANAALDEFAKQFRFGYDGHLLAEYORGVTIEASLTTISTALGAVSAGVSA 358
DB 326 IADKFRANKIEYSQRFKLGIDGDSLLAAAFHKTGAIDASLTTISTVLASVSSGISA 385
QY 359 AVGSAGVPTIALLVAGVTGLISGILEASKOAFESVANRLQKILEWKGQNGQYFKG 418
DB 386 ATTSILGAPVSALVAGVTGLISGILEASKOAFESVANRVKADVIAEWEKXH-KKNYFENG 444
QY 419 YDSRYAAYLANNLKFLSELNKELEAERVIAITQORWNNNIGELAGITKLGIERIKSKAYA 478
DB 445 YDARHAAPLEDFNFKLSQYNKEYSVSVSLVITQOHWDTLIGELAGVTRNGKDTLSGKSY 504
QY 479 DAFEDGKKVEAG-----SNITLDAKTGIIDISNSGKKTQALHFTSPLTAGTESRRLTN 534

```

Db 505 DYEEGRLEKKPDEFQKQVDPDLKGNIDLSDS--KSTLLKFTVPTLLTFCGEIRERRQS 562
Qy 535 KYSYINKLFRVKNQVTD-GEASSKLPFSKVIQVAETEGT-DEIGLIYNKAGNDD 592
Db 563 KKEYITELLVKGVDKWTVKGVQDKGSYDYSNLQHASVGNQYREIRIESHLGDDGDK 622
Qy 593 IFVGGKWNIDGGDHRVFSK--DGGFGNITVDTGTSATAGSVTVARKVARGD--IYHE 649
Db 623 VFLAAGSANIYAGKGDHVVYDKTDGTG--LTIDGTATKATEAGNTVTR-VLGGDVKVLQE 679
Qy 650 VVKRQETKVGKRTETIQYRDYELRKV-GYGYQSTDNLSKVEEVIGSQFNDVFKSGKFNDI 708
Db 680 VVKEQSVGKRTKTEQYRSYEFTHNGTDLTETDNLVSVEELIGTNRADKFFGSKFTDI 739
Qy 709 FHSSEGDLLDGGAGDRFLF----- 728
Db 740 FHGADGDHIBGNDGNRLYGDKNLTLRGNGDDQLYGGDNDKLTGGVGNNYLNGGDG 799
Qy 729 -----GGKGNDRLSGDEGDDLLDGGSDVLLGGAGNDVYIFPKGDGN 771
Db 800 DDELQVQNSLAKNVLSGKGNDRLYGSEGADLLDGGEGNDLLKGGYGNDIYRYSYGVH 859
Qy 772 DTLVDGTG-NDKLAFADANISDIMIERTKEGIIIVKRD-----HSGSINIPRWYITSN 823
Db 860 HIIDDDGGKDKLSLADIDFVAFKREGNDLIMYKAEGNVLSTGHKNGITFRNWF----- 915
Qy 824 LQNYQSNKTHKIEQLIGKQSGYTSQIDKIL---QDKDGTVITSOELKKLADENK-- 878
Db 916 -EKESGDISNHQIEQIPDKGRVITPDSLKKAPEYQQSNQANVYVEYASTYADLNLN 974
Qy 879 -----SQKLSAS--DI-----ASSLNKLVSMAFGTA-NSVSSNA 911
Db 975 PLINEISKIISAAGNFDVKEERSAASLLQSGNASDFSGRNSITLTA 1022

```

RESULT 26

```

Q8G924 PRELIMINARY; PRT; 1024 AA.
AC Q8G924
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE HlyA protein.
GN HlyA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=536;
RA MEDLINE=22267134; PubMed=12379716;
RA Dobrindt U., Blum-Oehler G., Nagy G., Schneider G., Johann A.,
RA Gotschalk G., Hacker J.;
RT "Genetic structure and distribution of four pathogenicity islands (PAI
RL Infect. Immun. 70:6365-6372(2002).
DR EMBL; AJ494981; CAD42039.1; -.
SQ SEQUENCE 1024 AA; 110365 MW; 3F3C284A45FEFB3 CRC64;

```

Query Match 39.5%; Score 1836.5; DB 2; Length 1024;
 Best Local Similarity 41.7%; Pred. No. 4.7e-74;
 Matches 420; Conservative 156; Mismatches 313; Indels 117; Gaps 24;

```

Qy 8 KSNIOAGLNSYKGLNLYLAIPKDYDPQKGTGLNDFDKADELGIARLAEEPNTETAK 67
Db 32 KDALKKAAEQTNRAGNRILLIIPKDYKQ-GSSLNDLVRTADELGIQVDEKNGTATK 90
Qy 68 KSDVTNVOFLSITGTGIAISATKLEKFLQKH--STNKLAKGLSDSVENIDRLKAGSNVLS 125
Db 91 QVFGTAELKIGLTERGVITFAPQLDKLQYKAGNKLKLG--GSAENIGDNLKAGSVLS 147

```

```

Qy 126 TSSFLGTALAGIELDSLKK-----GDAAPDALAKASIDLINEIIGNLSQSTQTIAPSS 181
Db 148 TFQNFGLTALSSMKIDELIKQKSGNVSSELAKASIELINQLVDTAAASLNNVNSFSQ 207
Qy 182 QLAKLSTTSIOAKGFNIGNKLQNL-NFSKTNIGLEIITGLLSGISGAFALADKNASTGK 240
Db 208 QLNKLASVLSNTHLNGVGNKLQNLPLNDINIGAGLDTVSGILSAISASFTLSNADADTGT 267
Qy 241 KVAAGPELSNOVIGNTKATISSVYLAQRAVAGLSTTGAVAAITSSIMLAISPLAFMNA 300
Db 268 KAAAGVELTIVKLVGNVKGISQYIIAQRAAQGLSTSAAGLSTAAVTLAISPLFSLIA 327
Qy 301 DKFNHALNDEFAQKPRKFGYDGDHLLAEYQORGVTIEASLTITISTALGAVSAGVSA 360
Db 328 DKFRANKIEYSORFKKLGDGDSLLAAPKGTGAIDASLTITISTVLASVSGISAAT 387
Qy 361 GSAVGTPIALLVAGVTGLISGLEASQAQMFESVANRLOQKILEWEKQNGQYFDKGYD 420
Db 388 TSLVGAPVSALVGAVTGIIISGLEASQAQMFESVANRLOQKILEWEKQNGQYFDKGYD 446
Qy 421 SRYAAVLANLKFELSELNKELEAERVIAITQQRWNNIGELAGITTKLGERIKSGKAYA 480
Db 447 ARHAAFLDEDFEILSQINKSEYSESVLITQQHWDTLIGELAGVTRNGDKTLSEKSYDY 506
Qy 481 FEDGKKVEAG-----SNITLDKGTGIIISNSNGKTKQALHFTSPLTAGTESRRLTNGK 536
Db 507 YESEKLEKEPDEFQKQVDPDLKGNIDLSVI--KSTLLKFTVPTLLTFCGEIRERRQS 564
Qy 537 YSYINKLKFGRVKNQVTD-GEASSKLPFSKVIQVAETEGT-DEIGLIYNKAGNDDIF 594
Db 565 YEYITELLVKGVDKWTVKGVQDKGSYDYSNLQHASVGNQYREIRIESHLGDDGDKV 624
Qy 595 VGQKWNIDGGDHRVFSK--DGGFGNITVDTGTSATAGSVTVARKVARGD--IYHEV 651
Db 625 LSAAGSANIYAGKGDHVVYDKTDGTG--LTIDGTATKATEAGNTVTR-VLGGDVKVLQE 681
Qy 652 KROETKVGKRTETIQYRDYELRKV-GYGYQSTDNLSKVEEVIGSQFNDVFKSGKFNDIF 710
Db 682 KEQSVGKRTKTEQYRSYEFTHNGTDLTETDNLVSVEELIGTNRADKFFGSKFTDI 741
Qy 711 SGEGDLLDGGAGDRFLF----- 728
Db 742 GADGDDHIEGNDGNRLYGDKNLTLRGNGDDQLYGGDNDKLTGGVGNNYLNGSGDGD 801
Qy 729 -----GGKGNDRLSGDEGDDLLDGGSDVLLGGAGNDVYIFRKGDNNDT 773
Db 802 ELQVQNSLAKNVLSGKGNDRLYGSEGADLLDGGEGNDLLKGGYGNDIYRYSYGVH 861
Qy 774 LYDGTG-NDKLAFADANISDIMIERTKEGIIIVKRD-----HSGSINIPRWYITSNLQ 825
Db 862 IDDDGGKDKLSLADIDFVAFKREGNDLIMYKAEGNVLSTGHKNGITFRNWF-----E 916
Qy 826 NYQSNKTHKIEQLIGKQSGYTSQIDKIL---QDKDGTVITSOELKKLADENK----- 878
Db 917 KESGDISNHQIEQIPDKGRVITPDSLKKAPEYQQSNQANVYVEYASTYADLNLNPL 976
Qy 879 -----SQKLSAS--DI-----ASSLNKLVSMAFGTA-NSVSSNA 911
Db 977 INEISKIISAAGNFDVKEERSAASLLQSGNASDFSGRNSITLTA 1022

```

RESULT 27

```

Q43892 PRELIMINARY; PRT; 1055 AA.
AC Q43892
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Leukotoxin.
GN Lkta.
OS Actinobacillus actinomycetemcomitans (Haemophilus
OC actinomycetemcomitans).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

```


OC Pasteurellaceae; Actinobacillus.
 OX NCBI_taxID=714;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JP2;
 RA Kolodrubetz J., Dailey T., Kraig E.;
 RT "Nucleotide sequence of the leukotoxin gene from actino-bacillus
 RT Actinomycetomcomtans: Homology to the A-hemolysin/leukotoxin gene
 RT family.";
 RL Infect. Immun. 58:920-929 (1991).
 DR EMBL; X16829; CAA34731.1; -
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR InterPro; IPR003995; RtxA.
 DR Pfam; PF00353; hemolysinCbind; 6.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PR00313; CABNDNGRPT.
 DR PRINTS; PR01488; RTXTXINA.
 DR PROSITE; PS00330; HEMOLYSIN CALCIUM; 1.
 SQ SEQUENCE 1055 AA; 113853 MW; 5331C396FA76669E CRC64;

Query Match 39.1%; Score 1816; DB 2; Length 1055;
 Best Local Similarity 40.2%; Pred. No. 4e-73;
 Matches 413; Conservative 185; Mismatches 299; Indels 130; Gaps 23;

QY 3 NINVIKSNIOAGLNTKSGLNLYLAIPKDYDPQKGGTLDNFKAADLGIAARLAEPNH 62
 DB 27 NIDAKEQLQKALDKLGTGKTKLTIYIPKNY--KKNGGLTALIKAAQKLGIEVHEGKG 84
 QY 63 TETAKSVDTVNOFLSLTGTGIAISATKLEKPLQ--KHSITNKLAKGLSDVENIDRKLGA 120
 DB 85 PALTINGILNTGKGLGLTGERLTLPAPELDKWIOGNKHLNSVG---STGNLTAKIDKV 140
 QY 121 SNVLSTLSFGLTALAGHELDSLIK---KGDAPDA-LAKASIDILNIEIIGLSOSTQTI 176
 DB 141 QSVLGLTQAFNTAFSGMDLALIKARQNGKNTVDQLAKASLNILNIEIGTITSSITNV 200
 QY 177 EAFSSOLAKLSTISOAGFSGNIGKNLQNL---NFSKTNLGLEITIGLSGISAFALA 232
 DB 201 DTFSQKLNKLGALGOVHFSGFGKLNLPKLGK---GLGALSGVLSAISALLLA 257
 QY 233 DKNASTGKKVAAGFELSQVIGNVTKAISVYLAQVAAGLSTTGAVAAALITSSIMLAIS 292
 DB 258 NKDAATATKAAAAEELTNKLVNIGKAITQYLAQRAAAGLSTTGAVAGLIASVVSLAIS 317
 QY 293 PLAFNNAADKFNHANALDEFAKQFKPGYDGHLLAEYVORGVTTEASLTITISTALGAVS 352
 DB 318 PLFLGLIAKQFARMLEYSKPKFKFGYNGSLGQFYKNTGIADAATITINTVLSAIA 377
 QY 353 AGVSAAGVSGAVGTPITALLVAGVTGLISGILEASQAMFESVANRLOQKILEWEKQNGQ 412
 DB 378 AGVGAASGSLVGAIPGLLVSAITSLISGILDASQAVFEHTANQALDKIKAWENKY-GK 436
 QY 413 NYFDKGYDSRYAAYLANNLKFLSELNKELEAEVIAITQQRWNNIGELAGITKLGRIK 472
 DB 437 NYFENGYDARHSAFLSLSKFLNELREKYKTNILSITQGWQDQRIAGLAGITRNGRIQ 496
 QY 473 SKGAYADAFEDKKVYAGSN----ITLDAKTGIIISNSNGKKTQALHFTSPLLTAGTES 528
 DB 497 SKGAYDVLKKEELAKHSDKFTKQLDPIKGNIDLSGKSTT--LTFNPLLTAGKEE 554
 QY 529 REELTNGKYSYINKLKFKGVKNQVTDGEASSKL-DESKVIQR-VAETEGTDEIGLVNA 586
 DB 555 RKTOSGKGYFEITELKVGKRTDVKGVNPNNGVNSGVDFSNLIQHAVTRDNKVLAEALIANL 614
 QY 587 KAGNDDIFVGGQKMNIDGGDRHDFVYSKDGFGNITVDGTSATBAGSYTVNRKVGARDI 646
 DB 615 GAKDYVFGSGSTIVNAGDGYDWDYSK-GRTGALTIDGRNATRAGQYKVERDLSGTVQ 673
 QY 647 YHEVVKRQTKVKKRTETITQYRDYEL--RKVGYGQSTDNLKSVEEVGSGQNFVFKGSK 704
 DB 674 LQETVSKQETKRGKVTDLLEYRNYKLDYIYNTKGFKAHDELNSVEEIIIGSTLDRDFYGS 733
 QY 705 FNDIFHSGSGDDL-----LOGGAGDDRLF----- 728

Db 734 FNDVFHGHGDDLLIYGYDGRDLRYGDNNDIEHGGQGNKLYGGAGNDRLFGYGNNYLD 793
 QY 729 -----GKGNDRLSGDEGDDLLDCGSGDDVNLGGAGNDVYIFRKGDG 770
 Db 794 GGEGLDHLGGNGSDILRGGSGDKLFGNQGGDDLLDGGEGDQLAGGEGNDIYYRKEYG 853
 QY 771 NDTLYDGTGN-DKLAFADANISDIMEIERTKEGIIIVKENDHSGSINIPRWYITSN---LQ 825
 Db 854 HHTITEHSGDKDKUSLANINLKDVSFERNGNDLLKTKNNRT-AVTFKGFKPKHSSAGLD 912
 QY 826 NYQ-----SNKTDHKEIQLEKDGSGVYTSQIDKILQDK 859
 Db 913 EYQRKLEAYAPEKDRARLKRQFELQRGVKDKSLNNKVEEIIIGKGERITSQIDNLFDKS 972
 QY 860 KDGVTITSQELKKL-ADENKSKQSASDIASS-----LNKLVGSNALFGTA-NS 906
 Db 973 GNKKTISPQELAGLIKNGKSSLSMSSSRSSMLTQKSGLSNDISRIISATSGFGSGKA 1032
 QY 907 VSSNALQ 913
 Db 1033 LSASPLQ 1039

RESULT 28
 O85101 PRELIMINARY; PRT; 998 AA.
 ID O85101;
 AC O85101;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hemolysin.
 GN EHXA.
 OS Escherichia coli.
 OG Plasmid EHEC-hemolysin plasmid.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_taxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=enterohemorrhagic EC920006;
 RX MEDLINE=98261495; PubMed=9596716;
 RA Boerlin P., Chen S., Colbourne J.K., Johnson R., De Grandis S.,
 RA Gyles C.;
 RT "Evolution of enterohemorrhagic escherichia coli hemolysin plasmids
 RT and the locus for enterocyte effacement in Shiga toxin-producing E.
 RT coli.";
 RL Infect. Immun. 66:2553-2561 (1998).
 DR EMBL; AF043471; AAC24352.1; -
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR InterPro; IPR003995; RtxA.
 DR Pfam; PF00353; hemolysinCbind; 6.
 DR Pfam; PF02382; RTX; 1.
 DR PRINTS; PR00313; CABNDNGRPT.
 DR PRINTS; PR01488; RTXTXINA.
 KW Plasmid.
 SQ SEQUENCE 998 AA; 107095 MW; E5A566B239DCBE55 CRC64;

Query Match 37.9%; Score 1762; DB 2; Length 998;
 Best Local Similarity 40.4%; Pred. No. 9.7e-71;
 Matches 396; Conservative 177; Mismatches 302; Indels 104; Gaps 24;

QY 8 KSNTOAGLNTKSGLNLYLAIPKDYDPQKGGTLDNFKAADLGIAARLAEPNHETAK 67
 Db 18 KSAFNATASSSVRSAGKLLIIPDNYEAQ-GVGINELVKAADLGIEIHRTRDDTAIAN 76
 QY 68 KSVDTVNOFLSLTGTGIAISATKLEKFLQKHSITNKLAKGL-DSVENIDRKLKASNVLST 126
 Db 77 QFFGTAEKVGLTERGVAFAPOLDKLLQKY--QKVGSKI GGTAENVGNLKGAGTVLSA 134
 QY 127 LSSFLGTALAGELDSLIKK----GDAAPDALAKASIDILNIEIIGLSQSTQTIAPFSQ 182
 Db 135 LQNFGTIALSGMALDELLRKOREGDISQNDIAKSSIELINQLVDTVSSINSTVDSFSQ 194


```
183 LAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLIITGLSGISAGFALADKNASTGK 241
195 LNQGSLSSKPRUSSVGGKLNQNPDLGPGDGLDVVSGILSASVAFILGNSDAHTGK 254
242 VAAGFELSNVIGNVTKAISSVYLAQRAAGLSTTGAAALITSSIMLAISPLAFMAAD 301
255 AAAGIELTTQVLGNVGRVSVQYILAQMAQGLSTTAASAGLITSAVLAISPLSFLAID 314
302 KFNHANNALDEPAKQFRKFGYDGHLLAERYQGVGTIEASLTITSTALCAVSAGVSAAVG 361
315 KFERAKQLESYSERFKFNTEYEGDALLAFAFHKEGAIDAAALTITNTVLSVSSAGVSAASA 374
362 SAVGTPITALVAGVTGLISGLEASKQAMFESVANRLQKLEWEKONGGNYFDKGYS 421
375 SLIGAPISMLVSALTGTISGLEASKQAMFHVADKFAARINEKEH-KGNYPENGIDA 433
422 RYAAAYLANNKFLSELNKELEAERVAITQQRWNNIGELAGITKGLERIKSGKAYADAF 481
434 RHAAFLEDSLSLLADFSRQHAVERAVAITQOHWEKIGELAGITRNADRSQSKAYINYL 493
482 EDGKKVQAG-----SNITLDAKTGLIIDSNSGKKTQALHTFSPLLTAGTSRRLTNGKY 537
494 ENGLLEAQPKFTQQVDFDPQKGTIDL--STGNVSSVLTFTVPTFTPEEVRERKQSGKY 551
538 SYINKLAKFGRVKNQWVTDGEASSK--LDFSKVIQORVAETEGTDEIGLIVNAKAGNDDI-F 594
552 EYMTSLIVNGKDTWSV-KGINKHKGVDYDYNLIQFV-EKDNKHQYQARIISLGDNDLIV 609
595 VGOQKMNIDGGDHRVYFSKDGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVVK 652
610 SGAGSSEVFAGEGHDTVSYNKT-DVGKLTIDATGASKPGEYIVSKNM-YGDVKVLQEVVK 667
653 ROETKVGKRTETIQRDYELRKVGYGYOSTDNLSKSEVIGSQNDVFKSGKFNDIFHSG 712
668 EQEVSVGKRTKIQYRDFEFRTGIPYDVIDNLHSHVEELIGCKHDDBFKGGKFNDIFHGA 727
713 -----EGDLDGCGAGDRLFGKGNDRLSGDEG----- 742
728 DGNDYIEGYNDRLYGDDGDDYISGGGQDDQLFGSGGNDKLSCGGDGNLYITGSGGNDL 787
743 -----DLDGSGDDVNLGAGNDVYIFRKDGNDLYD-G 777
788 QAHCAYNLSGCTGDDKLYGGGGIDLDGEGENDYINGFGNDIYVYRQNGVHTIADG 847
778 TGNDKLAPADANISDIMEIKETGIIVKR-----NDHSGSINIPRWYITSNLQN 826
848 GKGDRHLSDISFDDIAFKKVGNDLIMNKAINGALSFNESNDVNG-ITFKNWFADK---- 902
827 YQSNKTDHKTLEQLGKDSYITSQIDKILQDKDGTVITSQELKKLADENKSKLSASD 886
903 -ASGEDNHLVEVITDKGREIKADKISH-NNNEQSGYIKAS-----NTASE--KNMVNITS 954
887 IASSINKLVGSMALFGTAN 905
955 VANDINKIISVSGFDSGD 973

RESULT 29
Q9LC58 ID Q9LC58 PRELIMINARY; PRT; 998 AA.
AC Q9LC58,
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hemolysin A.
GN EHEC-HLYA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RA Taneike I., Wakisaka-Saito N., Harada Y., Zhang H.-M., Yamamoto T.;
RT "The enterohemorrhagic Escherichia coli (EHEC)-hemolysin genes of a
RT Shiga toxin 1 (Stx1)- and Stx2-producing, serotype O128 Escherichia
RT coli strain with a greatest hemolytic activity.";
RL Acta Med. Biol. (Nijigata) 0:0-0(2000).
DR EMBL; AB032930; BAA93708.1; -.
DR InterPro; IPR001343; Hemolysin_Ca_bind.
DR InterPro; IPR001395; RtxA.
DR Pfam; PF00353; hemolysinCbind; 6.
DR Pfam; PF02382; Rtx; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RXTXOXINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCITUM; 1.
SQ SEQUENCE 998 AA; 107196 MW; CD7A88E9BD862DB6 CRC64;

Query Match 37.9%; Score 1760; DB 2; Length 998;
Best local Similarity 40.4%; Pred. No. 1.2e-70;
Matches 396; Conservative 174; Mismatches 305; Indels 104; Gaps 24;

Qy 8 KSNIQAGLNSTKGLKNLYLAIPKDYDPQKGTGLNDFIKAADELGIARLAEEPNTETAK 67
Db 18 KSAFNATSSSVRAGKGLILLIPDNEYAQ-GVGINELVKAADDLGIHRTERRDDTAIAN 76
Qy 68 KSDVTNQFSLTQGTGTAISATKLEKFLQKHSNKLAKGL-DSVENIDRKLKGKASNYLST 126
Db 77 QFFGTAEKVVLGTERGVAIFAPQLDKLQKY--QKVGSKIGGTAEVNNLKGAGTVLSA 134
Qy 127 LSSFLGTALAGIELDSLKK-----GDAAAPALAKASIDLINELIGNLSQSTQTEAFSSQ 182
Db 135 LQNFPTGIALSGMALDELLRQREGEDISQNDIAKSSIELINQLVDTVSSINSTVDSFSEQ 194
Qy 183 LAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLIITGLSGISAGFALADKNASTGK 241
Db 195 LNQGSLSSKPRUSSVGGKLNQNPDLGPGDGLDVVSGILSASVAFILGNSDAHTGK 254
Qy 242 VAAGFELSNVIGNVTKAISSVYLAQRAAGLSTTGAAALITSSIMLAISPLAFMAAD 301
Db 255 AAAGIELTTQVLGNVGRVSVQYILAQMAQGLSTTAASAGLITSAVLAISPLSFLAID 314
Qy 302 KFNHANNALDEPAKQFRKFGYDGHLLAERYQGVGTIEASLTITSTALCAVSAGVSAAVG 361
Db 315 KFERAKQLESYSERFKFNTEYEGDALLAFAFHKEGAIDAAALTITNTVLSVSSAGVSAASA 374
Qy 362 SAVGTPITALVAGVTGLISGLEASKQAMFESVANRLQKLEWEKONGGNYFDKGYS 421
Db 375 SLIGAPISMLVSALTGTISGLEASKQAMFHVADKFAARINEKEH-KGNYPENGIDA 433
Qy 422 RYAAAYLANNKFLSELNKELEAERVAITQQRWNNIGELAGITKGLERIKSGKAYADAF 481
Db 434 RHAAFLEDSLSLLADFSRQHAVERAVAITQOHWEKIGELAGITRNADRSQSKAYINYL 493
Qy 482 EDGKKVQAG-----SNITLDAKTGLIIDSNSGKKTQALHTFSPLLTAGTSRRLTNGKY 537
Db 494 ENGLLEAQPKFTQQVDFDPQKGTIDL--STGNVSSVLTFTVPTFTPEEVRERKQSGKY 551
Qy 538 SYINKLAKFGRVKNQWVTDGEASSK--LDFSKVIQORVAETEGTDEIGLIVNAKAGNDDI-F 594
Db 552 EYMTSLIVNGKDTWSV-KGINKHKGVDYDYNLIQFV-EKDNKHQYQARIISLGDNDLIV 609
Qy 595 VGOQKMNIDGGDHRVYFSKDGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVVK 652
Db 610 SGAGSSEVFAGEGHDTVSYNKT-DVGKLTIDATGASKPGEYIVSKNM-YGDVKVLQEVVK 667
Qy 653 ROETKVGKRTETIQRDYELRKVGYGYOSTDNLSKSEVIGSQNDVFKSGKFNDIFHSG 712
Db 668 EQEVSVGKRTKIQYRDFEFRTGIPYDVIDNLHSHVEELIGCKHDDBFKGGKFNDIFHGA 727
Qy 713 EG-----DLDLDGCGAGDRLFGKGNDRLSGDEG----- 742
Db 728 DGNDYIEGYNDRLYGDDRDDYISGGGQDDQLFGSGGNDKLSCGGDGNLYITGSGGNDL 787
Qy 743 -----DLDGSGDDVNLGAGNDVYIFRKDGNDLYD-G 777
```

```
Db 788 QAHCAYNLTSGTGGDKLYGGGIDLLDGGEGNDYLGNGFGNGDIYVYQYNGHHTIADEG 847
Qy 778 TGNKDLAFADANISDIIMERTKEGIIYVR-----NDHSGSINIPRWYITSNLON 826
Db 848 GKGDRHLSDISFDIDIAFRVGNLIMNKANGVLSFNSNDVNG-ITFKWFAKD----- 902
Qy 827 YOSNKTDRHKBOLIGKSGSYITSDQIDKLODKKGTGVTITSOELKCLADENKSKLASD 886
Db 903 -ASGADNHLVEITDKQREIKADKIPH-NNNRSYIKAS-----NIASE--KMWVNITS 954
Qy 887 IASSLNKLVGSMALFGTAN 905
Db 955 VADDINKIISVSGFDSGD 973

RESULT 30
P71223
ID P71223 PRELIMINARY; PRT; 998 AA.
AC P71223;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE EHEC-hemolysin.
GN EHEC-HLYA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=78-92;
RX MEDLINE=97034824; PubMed=8880480;
RA Schmidt H., Karch H.;
RT "Enterohemolytic phenotypes and genotypes of Shiga Toxin-producing
RT Escherichia coli O111 strain from patients with diarrhea and hemolytic-
RT uremic syndrome."
RL J. Clin. Microbiol. 34:2364-2367 (1996).
DR EMBL; X94129; CA63849.1; -.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCabin; 6.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
SQ SEQUENCE 998 AA; 10725 MW; DA3BF078C7E4131E CRC64;

Query Match 37.8%; Score 1756; DB 2; Length 998;
Best Local Similarity 40.1%; Pred. No. 1.8e-70;
Matches 393; Conservative 176; Mismatches 306; Indels 104; Gaps 21;

Qy 8 KSNIQAGLNSTKSGLNKLYLAIPKDYDPQKGTGLNDFKAADELGIARLAEBPNHTETAK 67
Db 18 KSAFNATSSVRSAGKLLIPDPNVEAQ-GVGINELVKADELGIHRTERDDTAIAN 76
Qy 68 KSDVTNQPLSTQTGIAISATKLEFKLQKHT--NKLAKGLDSVENIDRLKGNVLS 125
Db 77 QPFGAAEKVVGUTERGVAIFAPQLQKQVGSIGR----TAENVNNGKAGTVLS 133
Qy 126 TLSSFLGTALAGIELDSLTKK----GDAAAPDALAKASIDLINEIIGLSQSTQTIEAFSS 181
Db 134 ALQNTGTIALSGWALDELARKQREGEDISQNDIAKSSIELINQLVDTVTSSINSTVDSFE 193
Qy 182 QLAUKGSTISQAKGFSNIGNKLNLFNFKTNLGLIEITGLLSGISAGFALADKNASTGK 240
Db 194 QLNQLGSLFSLSPRLSSVGGKLNLPDLGLSLGDLVDVWSGILSAVSASFILNSDAHTGT 253
Qy 241 KVAAGFELSNQVIGNVTKAISSVLAORVAAGLSTTGAAVAILTSIMLAISPLAFMNA 300
Db 254 KAAAGIELTQVIGNGVKASQVILQORVAQGLSTTAASAGLITSAVMLAISPLFLAA 313
Qy 301 DKFNHANALDEFAPKRFKFGDGHLLAEYQRGVGTIEASLTITSTALGAVSAGVSAAV 360
Db 314 DKFERAKQLESYERFKLNYEGDALLAAPHKETGAIDAALTINTVLSSVSAGVSAASS 373
```

```
Qy 361 GSAVGTPIALLVAGVTGLISGLEASQAMFESVANRLOQKILEWEKQNGQNYFDKGYD 420
Db 374 ASLIGAPISMLVSLTGTISGLEASQAMFEHVAEKFAARINWEKEH-GKYNFENGYD 432
Qy 421 SRVAAAYLANLKFSELNKELEAEVTAITQORWNNNIGELAGITKIGERIKSKAYADA 480
Db 433 ARHAAPLEDLSLLADFSRQHAVERAVAITQOHWEKIGELAGITRNADRSQSKAYIN 492
Qy 481 FEDGKKVEAG-----SNITLDAKTGIIDISNNGKKTQALHFTSPLLTAGTRESRLTNGK 536
Db 493 LENGGLLEAQPKEFTQQVFPQKGTIDL--STGNVSSVLTFTFTFTFGBEVRERKOSGK 550
Qy 537 YSYINKLKFRGVKNQVTDGEASSK--LDPSKVIQORVAETEGTDEIGLIIVNAKAGNDIF 594
Db 551 YEYNTFLVNGKDTWSV-KGIKNHKGYDYSKLIQFVEKTKYQARMSLGDKDDVY 609
Qy 595 VGQKMNIDGGDHRVYFYSKGGFGNITVDGTSATEAGSYTVNRKVARGD--TYHEVVK 652
Db 610 SGAGSSEVFAGEGHDVPVYNKT-DVGKLTIDATGASKPGEYIVPKNM-YGDVEVLQEVVK 667
Qy 653 ROETKVGKRTETIQRDYELRKVGYGQSTDNLSVEEVIGSQFNDVFKSGKFNDIPIHSG 712
Db 668 EQEVSVGKRTKIQYRDFEFTGTGIPYDVIDNLHSHVBEELIGGKHDDFKGKGFNDIPIHGA 727
Qy 713 -----EGDLDLGGAGDRLFGGKGNDRLSGDEG----- 742
Db 728 DGNDYIEGNYGNDRLYGDDGDDYISGGQDDQLFEGSGNDKLSGGDGNVYITGSGNDEL 787
Qy 743 -----DLLDGGSGDDVLANGAGNDVYIFRKQDGNLTLYD-G 777
Db 788 QAHCAYNLTSGTGGDKLYGGGIDLLDGGEGNDYLGNGFGNGDIYVYQYNGHHTIADEG 847
Qy 778 TGNKDLAFADANISDIIMERTKEGIIYVR-----NDHSGSINIPRWYITSNLON 826
Db 848 GKGDRHLSDISFDIDIAFRVGNLIMNKANGVLSFNSNDVNG-ITFKWFAKD----- 902
Qy 827 YOSNKTDRHKBOLIGKSGSYITSDQIDKLODKKGTGVTITSOELKCLADENKSKLASD 886
Db 903 -ASGADNHLVEITDKQREI--KVDKIPHNNE----RSGYIKASNIASEKSVNITS 954
Qy 887 IASSLNKLVGSMALFGTAN 905
Db 955 VADDINKIISVSGFDSGD 973

RESULT 31
Q46716 PRELIMINARY; PRT; 998 AA.
AC Q46716;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hemolysin A (HLYA).
GN HLYA OR EHEC-HLYA.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=98290540; PubMed=9628576;
RA Makino K., Ishii K., Yasunaga T., Hattori M., Yokoyama K.,
RA Yaeudo H.C., Kubota Y., Yamaiichi Y., Iida T., Yamamoto K., Honda T.,
RA Han C., Ohtsubo A., Kasamatsu M., Hayashi T., Kuhara S.,
RA Shinagawa H.;
RT "Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
RT enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
RT outbreak."
RL DNA Res. 5:1-9 (1998).
RN (2)
RP SEQUENCE OF 241-998 FROM N.A.
```



```

Db 258 DADACTKAAAGIEISTKILGNIGKAVSOYIIIAQORVAAGLSITTAATRWFNRSVVALAISPL 317
Qy 295 AFMNAADKFNHANAALDEFAKFRKFGYDGDHLLABYQYRGVGTIBASLTITITAGLAVSAG 354
Db 318 SPLNADVAFERAKQLEQYSEKFKFGYEGDLSLASFYRETGAIEAALTTINSVLISAAPAG 377
Qy 355 VSAARAVGAVGTPIALLVAGVTGLISGLEASKQAMFESVANRLOGKILEWKQNGGONY 414
Db 378 VGAAATGSLVGAFAVAALSAITGIISGLIDASKQAFERVAITKLANKIDEWKXH-GKNY 436
Qy 415 FDKGDSYAAAYLANNKFLSELNKELEAERVAITQORWNNIGELAGITKLAGERIKSG 474
Db 437 FENGYDARHSAFLEDTFELLSONKEYSVERVAITQORWNNIGELAGITKLAGERIKSG 496
Qy 475 KAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSGKKTQALHFTSPILLTAGTERE 530
Db 497 KAYVDFELEGKLEKDPDFKVPDLEGKIDLSIN--KTTLKLTFTPAFTAGEETRE 554
Qy 531 RLTKGKYSYINKLKFRGVKNQVTDGEASSKL-DFSKVIQORVAETEGTDEGLIIVNAKAG 589
Db 555 RKQTKYEITWELTFVKGKEKVVVTGVQSHNAILYDINLIQALDKKG-EKQVITIESHLG 613
Qy 590 --NDDIFVQGGKWNIDGGDGHDRVFYSK-DGGFGNITVDGTSATBAGSYTNNRVY-ARGD 645
Db 614 EKNDRIYLSGSSSIEVAGNGHDVAYYDKDTGY--LTFDGSQAQAGYIVTKELKADVK 671
Qy 646 IYHEVVKQETKVGKRTETIQYRDYELR--KVGYGYQSTDNLSKVEEVIGSQFNDVFKGS 703
Db 672 VLKEVVKQDQISVGRSEKLEYRDELSPFELNGIRAKDELHLSVEEIGSNRDKPFGS 731
Qy 704 KENDIFHSGEGD-----DLDGAGGDDRLFGKGNDRLSGDEGDDL 745
Db 732 RTFDIFHGAAGDDEIYNGDHDILYGGDNDVIHGGDGNHLLVGGNDRILIGKGNPL 791
Qy 746 DGGSGDD-----VLNGAGNDVY-----IFRKGNDPLYDGTGND----- 781
Db 792 NGGDGDELQVPECQYVNLGAGNDILYSGDGNLFDGGVNDKIYGLGKDIYRSKE 851
Qy 782 -----KIAFADANI-----SDIMIBRTKEGIIIVKRNDSHG-SINIPR 817
Db 852 YGRHIIIEKGGDDTLALLSDLSFKDVGFTIRIGDILLVKNRIGTGLYHYHEDYNGNALTIKD 911
Qy 818 WIYTSNLQYQSNKTDHKEQLIGKGSVITSQDKILQDKDGTVITSBELKLAEN 877
Db 912 WF--KEGEGQNN-----KVEKIVDKGAYVLSQYLTETAPGRNGYNGLEEKLYYEG 965
Qy 878 KSQ-----KLSASDIASSNLKLVGS-----MALFGTANSVSSNALOPITQ 917
Db 966 YNALPQLRKDIBQIISSTGALTGEHQVLVGAGGPLAYSNSPNSIPNAPSNYLTQ 1020

RESULT 33
Q47461
ID Q47461 PRELIMINARY; PRT; 998 AA.
AC Q47461;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE EHEC-hlyA protein.
OS EHEC-HLYA.
OS Escherichia coli.
OS Plasmid p0157.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EDL 933;
RX MEDLINE=95172699; PubMed=7868227;
RA Schmidt H., Beutin L., Karch H.;
RT "Molecular analysis of the plasmid-encoded hemolysin of Escherichia
col 0157:H7 strain Edl933.";
RT Infect. Immun. 63:1055-1061(1995).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=EDL 933;
RX MEDLINE=97090409; PubMed=8936317;
RA Schmidt H., Kernbach C., Karch H.;
RT "Analysis of the EHEC hly operon and its location in the physical map
of the large plasmid of enterohemorrhagic Escherichia coli 0157:H7.";
RT Microbiology 142:907-914(1996).
DR EMBL; X86087; CAAG0042.1; -.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003995; RTX.
DR Pfam; PF02382; hemolysinCabin; 6.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CABNDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
KW Plasmid.
SQ SEQUENCE 998 AA; 107032 MW; 4D8BE108C309BF7E CRC64;

Query Match 37.3%; Score 1734; DB 2; Length 998;
Best Local Similarity 39.9%; Pred. No. 1.7e-69;
Matches 390; Conservative 174; Mismatches 312; Indels 102; Gaps 21;

Qy 8 KSNTOAGLNSKSGKLNLYLAIPKDYDPQKGTLDNFIKADELGIARLAEPEHHTETAK 67
Db 18 KSAFTASSSVRSAGKKLILLIPDNYEAQ-GVGINELVKADELGIEIHRTERDDTAIAN 76
Qy 68 KSVDTVNOFSLTGTGTAISATKLEKLEKLOKHSNTKLAKGL-DSVENIDRKLKGSANVLST 126
Db 77 QFFGAEEKVGLTERGVAIFAPQUDKLLQKY--QKVGSKIGGTAEENVGNNGNKGAGTVLSA 134
Qy 127 LSSFLGTALAGIEDLSLKK----GDAAPDAKASIDLNEIICNLSQSQTIEAFSSQ 182
Db 135 LQNTGTALSGMALDELLRKQREGEDISQNDIAKSSIELINQVDTVSSINSTVDSFSEQ 194
Qy 183 LAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLIEIITGLSGISAGPALADKNASTGKK 241
Db 195 LNQLGSFLSSKPRSSVGGKLNQLPDLGLDGLDVVSGILSAVSASFILGNSDAHTGK 254
Qy 242 VAAGFELSNQVIGNVTAISSYVLAORVAAGLSITGAVAALITSSIMLAISPLAFMNAAD 301
Db 255 AAAGIELTQVLGNVGVKAVSQYIIAQRMAQGLSTTAASAGLITSAVMLAISPLSFLAAD 314
Qy 302 KFNHANALDEFAKFRKFGYDGDHLLABYQYRGVGTIBASLTITITAGLAVSAGVSAAVG 361
Db 315 KFERAKQLESYSEKFKLYEGDALLAGFKETGAIDAGLTINTVLSVSAVSAASA 374
Qy 362 SAVGTPIALLVAGVTGLISGLEASKQAMFESVANRLOGKILEWKQNGGONYFDKGVDS 421
Db 375 SLICAPISMLVSALTGTISGILEASKQAMFESVAAKFAARINEWEKEH-KKNYFENGYDA 433
Qy 422 RYAAAYLANNKFLSELNKELEAERVAITQORWNNIGELAGITKLAGERIKSGKAYADAF 481
Db 434 RHAAFLPDESLDLSLADFSRQHAVERAVAITHOHWDEKIGELAGITRNADRSQSGKAYINYL 493
Qy 482 EDGKKVEAG----SNITLDAKTGIIDISNSGKKTQALHFTSPILLTAGTERERLTNGKY 537
Db 494 ENGGILEAQPEFTQOVDFPKGTIDL--STGNVSSVLTFITPTFPEEVRERKQSGKY 551
Qy 538 SYINKLKFGRVKNQVTDGEASSK--LDFSQVIRVAETEGTDEGLIIVNAKAGNDDIFV 595
Db 552 EYMTSLIVNGKDTWSV-KGIKNHKGVDYDKLIQFVEKNTKHYQARIISELGKDDVVYS 610
Qy 596 GQGNKNDIGDGHDRVFYSKDDGGFGNTIVDGTSTATEAGSYTVNRKVARGD--IYHEVVKR 653
Db 611 GAGSEVFAGEGYDPVSYNKT-DVGKLTIDATGAPKPGEYIVPKNM-YGDEVILQEVVKE 668
Qy 654 QETKVGKRTETIQYRDYELRKVGYGQSTDNLSKVEEVIGSQFNDVFKGSKFNDIFHSG- 712
Db 669 QEVSVGRKTEKIQYRDFEFTGGIPYDVIDNLHSVEELIGGKHDDDEFKGGKFNDFHGD 728
Qy 713 -----EGDLDLGGAGDRLFGKGNDRLSGDEG----- 742
Db 729 GNDYTEGYNDRLYGDDDDYISGGQGDQLFGSGGNDKLSGGDGNNGYLTGGSGNDBLQ 788

```


| RA | Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A., Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R., Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A., Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kicajima J.P., Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C., Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A., Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L., Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y., Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B., Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S., Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A., de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A., Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B., Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M., de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E., da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr., da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A., de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchiko M.H., Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L., Zago M.A., Zatz M., Meidanis J., Setubal J.C.; | 25 | LYLAIPKDYDPQKGGTLNDFIKAADELG--- | 793 | MYLALP--- | 82 | TG--- | 842 | NGVMAMDASGLEAKLDQTVHKNKAQALQDVMILYR--- | 140 | --- | 890 | PFDAIRHMDRTAATPPGQALAAAEITLV--- | 193 | AKGFSNTGNKLQNLNFKSTNLGLBEITG--- | 943 | GGG--- | 248 | LSNQVIGNVTKAI--- | 982 | GSNRLEGAGNDVLKVSYSADNVLIG--- | 306 | ANALDEPAKFRKFG--- | 1033 | DTTIEQGGTKLVPEGEGILAADVRLREGQDVLDLNGHDSIRLKDWLTSNGTRNSHAD | 341 | LTITISTALGAVSAGVSAAVG- | 1093 | IEQIVFADGTLWTPETLSSMGLTTLTGTPGNDTLKGWQKDILLGGAGDDVLDGEG--- | 393 | SVANRLOG--- | 1149 | --SNRLEGAGDDVLKVSYSADNVLISGG--- | 443 | AERVIAITQORWNNIGELAGITKL--- | 1193 | --- | 501 | IIDISNSNGKKTQALHFTSPLLTAGTESRRRLTNGKYSYINKLKFGVQWQVTDGEASS |
|----|---|----|---------------------------------|-----|-----------|----|-------|-----|--|-----|-----|-----|---------------------------------|-----|---------------------------------|-----|--------|-----|------------------|-----|------------------------------|-----|-------------------|------|---|-----|------------------------|------|--|-----|-------------|------|---------------------------------|-----|-----------------------------|------|-----|-----|--|
| RA | Frage J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R., Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A., Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kicajima J.P., Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C., Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A., Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L., Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y., Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B., Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S., Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A., de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A., Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B., Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M., de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E., da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr., da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A., de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchiko M.H., Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L., Zago M.A., Zatz M., Meidanis J., Setubal J.C.; | 25 | LYLAIPKDYDPQKGGTLNDFIKAADELG--- | 793 | MYLALP--- | 82 | TG--- | 842 | NGVMAMDASGLEAKLDQTVHKNKAQALQDVMILYR--- | 140 | --- | 890 | PFDAIRHMDRTAATPPGQALAAAEITLV--- | 193 | AKGFSNTGNKLQNLNFKSTNLGLBEITG--- | 943 | GGG--- | 248 | LSNQVIGNVTKAI--- | 982 | GSNRLEGAGNDVLKVSYSADNVLIG--- | 306 | ANALDEPAKFRKFG--- | 1033 | DTTIEQGGTKLVPEGEGILAADVRLREGQDVLDLNGHDSIRLKDWLTSNGTRNSHAD | 341 | LTITISTALGAVSAGVSAAVG- | 1093 | IEQIVFADGTLWTPETLSSMGLTTLTGTPGNDTLKGWQKDILLGGAGDDVLDGEG--- | 393 | SVANRLOG--- | 1149 | --SNRLEGAGDDVLKVSYSADNVLISGG--- | 443 | AERVIAITQORWNNIGELAGITKL--- | 1193 | --- | 501 | IIDISNSNGKKTQALHFTSPLLTAGTESRRRLTNGKYSYINKLKFGVQWQVTDGEASS |
| RA | Frage J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R., Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A., Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kicajima J.P., Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C., Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A., Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L., Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y., Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B., Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S., Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A., de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A., Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B., Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M., de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E., da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr., da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A., de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchiko M.H., Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L., Zago M.A., Zatz M., Meidanis J., Setubal J.C.; | 25 | LYLAIPKDYDPQKGGTLNDFIKAADELG--- | 793 | MYLALP--- | 82 | TG--- | 842 | NGVMAMDASGLEAKLDQTVHKNKAQALQDVMILYR--- | 140 | --- | 890 | PFDAIRHMDRTAATPPGQALAAAEITLV--- | 193 | AKGFSNTGNKLQNLNFKSTNLGLBEITG--- | 943 | GGG--- | 248 | LSNQVIGNVTKAI--- | 982 | GSNRLEGAGNDVLKVSYSADNVLIG--- | 306 | ANALDEPAKFRKFG--- | 1033 | DTTIEQGGTKLVPEGEGILAADVRLREGQDVLDLNGHDSIRLKDWLTSNGTRNSHAD | 341 | LTITISTALGAVSAGVSAAVG- | 1093 | IEQIVFADGTLWTPETLSSMGLTTLTGTPGNDTLKGWQKDILLGGAGDDVLDGEG--- | 393 | SVANRLOG--- | 1149 | --SNRLEGAGDDVLKVSYSADNVLISGG--- | 443 | AERVIAITQORWNNIGELAGITKL--- | 1193 | --- | 501 | IIDISNSNGKKTQALHFTSPLLTAGTESRRRLTNGKYSYINKLKFGVQWQVTDGEASS |
| RA | Frage J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R., Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A., Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kicajima J.P., Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C., Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A., Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L., Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y., Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B., Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S., Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A., de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A., Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B., Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M., de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E., da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr., da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A., de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchiko M.H., Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L., Zago M.A., Zatz M., Meidanis J., Setubal J.C.; | 25 | LYLAIPKDYDPQKGGTLNDFIKAADELG--- | 793 | MYLALP--- | 82 | TG--- | 842 | NGVMAMDASGLEAKLDQTVHKNKAQALQDVMILYR--- | 140 | --- | 890 | PFDAIRHMDRTAATPPGQALAAAEITLV--- | 193 | AKGFSNTGNKLQNLNFKSTNLGLBEITG--- | 943 | GGG--- | 248 | LSNQVIGNVTKAI--- | 982 | GSNRLEGAGNDVLKVSYSADNVLIG--- | 306 | ANALDEPAKFRKFG--- | 1033 | DTTIEQGGTKLVPEGEGILAADVRLREGQDVLDLNGHDSIRLKDWLTSNGTRNSHAD | 341 | LTITISTALGAVSAGVSAAVG- | 1093 | IEQIVFADGTLWTPETLSSMGLTTLTGTPGNDTLKGWQKDILLGGAGDDVLDGEG--- | 393 | SVANRLOG--- | 1149 | --SNRLEGAGDDVLKVSYSADNVLISGG--- | 443 | AERVIAITQORWNNIGELAGITKL--- | 1193 | --- | 501 | IIDISNSNGKKTQALHFTSPLLTAGTESRRRLTNGKYSYINKLKFGVQWQVTDGEASS |
| RA | Frage J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R., Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A., Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kicajima J.P., Krieger J.E., Kurama | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |


```
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.; Xylella fastidiosa.";
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL; AE003910; AAF83478.1; -.
DR InterPro; IPR001343; HemlyenCa bind.
DR Pfam; PF00353; hemolysinCabin; 15.
DR PRINTS; PR00313; CAENDNGRPT.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 3.
DR Complete proteome.
SQ SEQUENCE 1208 AA; 128403 MW; DF151ACD7741872B CRC64;

Query Match      8.4%; Score 411; DB 16; Length 1208;
Best Local Similarity 22.0%; Pred. No. 3.1e-10;
Matches 227; Conservative 169; Mismatches 344; Indels 294; Gaps 46;

QY 36 QKGGTLNDFIKAADELGIARLAEPNHTTAKSVDTVNQFL-----SLTGTG-----IA 85
DB 286 QDQSQANLEFLTSSLGIASITLPTNTE---DVLGNGLNDRGYTRTDGRTGVVG 341
QY 86 ISATKLEKFLQKHS-----TNKLAGLDSVENID--RKLKASNVLSLSSFLGTALA 136
DB 342 DLQLGLEHYFDYSGAHDKVTVDAAALPAISGSAVRDLREASLSPALLAAVQALTP 401
QY 137 GIELSLIKKGDAAPDALAKASIDLINEIIGNLS-----170
DB 402 GTTRDTM-----RAALDPLVALWAGTSAMPSTEQRLTSAAPRVVYHGA VP 449
QY 171 -----QSTQTTEAF--SSQAKLGISTISQAGFSNIGNKLNLFNKTNLGLEIITGLLSG 224
DB 450 AAVTAQGEAVQAWIQOQHARLGPPIIAILEKFN--GSSLVSRNGQVSTGETFTW---- 503
QY 225 ISAGFALADKNASTGKVAAGFELSNOVI---GNVTKAISSVLVAQRAAGLSTTGAVAA 281
DB 504 -----NRVTHPDGSHSDYMSLLOPEQINSLMSAYASLUKE 538
QY 282 LITSSIMLA-----ISPLAFNNAADKFN--HANALDEFAPKPRGYDGDHLLAE---Y 330
DB 539 FAYAGLVLSLRLSDVLSGLTLYSNDAGWDASALE--AKLDHTWQHNKTQALQDVMDLY 596
QY 331 QRGVGTIEASLTFTTISTALGAVSAGVSAAGVAGTPIALIVAGVTGLISGILEASKQA- 389
DB 597 RYGSNAVAASGKPPFDALRHMDRTAATPDGRQ-----ALAAAEIT-LVSGSAEGSTAAD 650
QY 390 ---MPSVANRLOKLEWEKQNGQN-----YFPKGYDYSYAYLANNLKF 434
DB 651 LIFGDAGANILRGAGD--DVLSSGSDGDDTLYGCGDNDTLYGAGNDTLTGG--GGNDLLL 707
QY 435 -----SELNKELE---AERVIAITQQRWNNIGELAGITKLGRIKSGKAYADAF 481
DB 708 GGGDDVDLGGGSRNLEGGAGDDVLKVSYSWADNVL-----IGTGDDTLGSAFADTY 762
QY 482 ----EDGKK--VEAGSNITLDAKTGII--DISNSNGKKTQALHFTSPILLTAGTESRERLT 533
DB 763 LFNQGDGHDHTIEQGGTKLVFEGEIVAADVR-----LLRQGDQVVDLIG 807
QY 534 NKYSYINKLKGPRVKNVQVDTGEASSKLDFSKVTRVAETECT-----DEIGLIYNA 586
DB 808 NGHDSI-----RLKDWLTSNGTRNHSAD-----IEQIVFADGTTLTPTLSSMGLTTIG 856
QY 587 KAGND-----DIFV--GQGRKNNIDGGDHRV-----FYSKD-----GFGFN 621
DB 857 TSGNDTLKQWQKQDILLGAGDDVDLGGEGSRNLEGGAGNDVLKVSYSWADNVLIGTGTD 916
QY 622 ITVDGTSATPAGSYTVNRKVRAGDIYHEVVKROET-----656
DB 917 DTLGSAF--ADTYLFN---QGDGHDHTIEQGGTKLVFEGEIVAADVRLIREGQDVVL 970
QY 657 KVKGKTETIQRDYELRKVGYGYQSTDNLKSVEVI-----GSGQF 696
DB 971 DLGNHGDHSIRLKDW--LTSNGTRNHSAD-----IEQIVFADGTTLTPTLSSMGLTTIGTPG 1025
QY 697 NDVFKGSKFNDIFHSGEGDLDLGGAGDRLFGKGND--RLSGDEGDDLLDGGSGDDVL 754

Db 1026 NDTLKGWQKIDLLGGDDVDLGGEGSRNLEGGAGNDVLKVSYSWADNVLIGTGDDTL 1085
QY 755 NGAGNDVYIFRKGNDNTLYDGTGNDKLFADANA--NISDIMIERTEKGIIVKRNHSGSI 813
DB 1086 YGSAFADTYLFNKGDDGHDHTIEQSGTKLVFEGEGLHOKEARFTKSGDGLSLFNGSEDQV 1145
QY 814 NIPRWYITSNLQVNSNKTDKHIEQLIKGDKGSYITSDIKILQDKDKDGTVTTSOELKKL 873
DB 1146 TIAGWF-----NGSGHQVESLVFDGT--VLSGEVERLIAAMALSPPAATTLQ--ASV 1193
QY 874 ADENKSKQLSASDI 887
DB 1194 GDHKESPRLVASSI 1207

RESULT 39
Q9PAT8 PRELIMINARY; PRT; 2064 AA.
AC Q9PAT8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Bacteriocin.
GN XF2407.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
RX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
EX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.B., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Pacinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohlseil J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.B.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshukako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL; AE004049; AAF85206.1; -.
DR InterPro; IPR001343; HemlyenCa bind.
DR Pfam; PF00353; hemolysinCabin; 31.
DR PRINTS; PR00313; CAENDNGRPT.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 7.
KW Complete proteome.
SQ SEQUENCE 2064 AA; 218776 MW; 6265655E90DA1D9A CRC64;

Query Match      8.4%; Score 391; DB 16; Length 2064;
Best Local Similarity 25.4%; Pred. No. 5.2e-09;
Matches 135; Conservative 89; Mismatches 163; Indels 144; Gaps 23;
```

| | | | |
|----|------|---|------|
| Qy | 446 | VIATQORWNNIGELAGITKLERIKSGAYADAF-----EDGKK--VEAGSNITLDAKT | 499 |
| Db | 1588 | VLKAYRSADNVL-----IGTGDDTLTLYGSAYADTYLFNKGDDHDTTIEQGGDDTLVFGA | 1642 |
| Qy | 500 | GIIDISNSNGKTKQALLHFTSPLLTAGTYESRERLTNGKYSINKLKFGRVKRWQVTDGEAS | 559 |
| Db | 1643 | GIV-----ASQV-----RVLREGQDVVLDLGNGHDSI-----RLKDWLTSDGYRN | 1682 |
| Qy | 560 | SKLDFSKVIQORVABTEGT-----DEGLIVNAKAGNDDIFVQOGKMNIDGGDGHDRVF | 612 |
| Db | 1683 | GNND-----IEQIVFADGTITWTPETLSMGUJTLTSGNDTLKGWQGXDIILGCGDDVVL-- | 1737 |
| Qy | 613 | YSKDGFGF-----NITVDGT-----SATEAGSYTVNRKVARGD | 645 |
| Db | 1738 | ---DGGWGSNRLEGGAGNDVLKVAYRSADNVLTGTTGDDTLYSGYADTYLFN-----KGD | 1790 |
| Qy | 646 | IYHEVVKR--OETKV-----GKRTETIQVRVYELRKVGYGYQ | 680 |
| Db | 1791 | GHDITIIIEQGGDDTLVFGAGIVASQVRVLRGQDVVLDLGNHDSIRLKW-----LTSQYR | 1847 |
| Qy | 681 | STDNLKSVBEVI-----GSOFNDFVFKSGKFNDIPHSGEGDDLDDG | 720 |
| Db | 1848 | NGNN--DIEQLIVFADGTITWTPETLSMGLTTLTGTSNGDTLKGWQXDIILGCGDDVLDG | 1905 |
| Qy | 721 | GAGDDRLFGKGKND--RLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFPKGDXGNDTLVDGT | 778 |
| Db | 1906 | GMGSNRLEGGAGNDVLKVAYRSADNVLTGTTGDDTLYGSAYADTYLFNKGDDHDTTIEQG | 1965 |
| Qy | 779 | GNDKLAFADANISIMIERTEKG--IIVKRNHDSGSIINPRWYITSNLQNVQSNKTDHKI | 836 |
| Db | 1966 | GDDTLVFG-AGLHQKEARFTHSGNDLSILFNASEDQVTIAGWF-----NGSQHV | 2014 |
| Qy | 837 | EQLIGKDSYITSDQIDKILQDKDGTVITSQELKKLADENKSKLSASDI | 887 |
| Db | 2015 | ESLIVFDGCT-VLSGEVERLIAAMALSAVITTO-ASVRDTKESHRLVASSI | 2063 |

RESULT 40

| | | | |
|--------|--|------|---------|
| Q51868 | PRELIMINARY; | PRT; | 208 AA. |
| ID | Q51868 | | |
| AC | Q51866; | | |
| DT | 01-NOV-1996 (TrEMBLrel. 01, Created) | | |
| DT | 01-NOV-1996 (TrEMBLrel. 01, Last sequence update) | | |
| DT | 01-MAR-2003 (TrEMBLrel. 23, Last annotation update) | | |
| DE | Leukotoxin A (Fragment). | | |
| OS | Pasteurella haemolytica. | | |
| OC | Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; | | |
| OC | Pasteurellaceae; Mannheimia. | | |
| OX | NCBI_taxid=75985; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=T4 SEROTYPE AND T15 SEROTYPE; | | |
| RA | Lainson A.F., Aitchison K.D., Donachie W.; | | |
| RL | Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases. | | |
| DR | EMBL; Z22885; CA80499.1; -. | | |
| DR | EMBL; Z22886; CA80500.1; -. | | |
| DR | HSSP; Q03023; IAKL. | | |
| DR | InterPro; IPR001343; Hemlysen_Ca_bind. | | |
| DR | Pfam; PF00353; hemolysinCabin; 3. | | |
| FT | NON TER | | |
| SO | SEQUENCE 208 AA; 22803 MW; 25301410C85A4CC5 CRC64; | | |

| | | | |
|----|-----|--|-----|
| Qy | 839 | LIHGSGVITTSDOJDKILOKDKGTVTTSQBLKKLADENKSOKLSASDTASSINKLAVGSM | 839 |
| Db | 119 | IIQNGERITTSQVDELKEKG--KIDKSDLSUQVVDNYQLKYS--RDASNSLDKLISSA | 175 |
| Qy | 899 | ALFTGANSVSSNALQPIQTPTQGILAPSV | 927 |
| Db | 176 | SAFTSSND--SRNVL---ASPT--SMLDPSSL | 199 |

Search completed: February 17, 2004, 10:11:03
Job time : 57 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:18:40 ; Search time 21 Seconds

(without alignments)
1867.722 Million cell updates/sec

Title: US-10-069-799-5

Perfect score: 927

Sequence: 1 MSNINVIKNIQAGLNSTKS.....SSNALQPIPTQGTGILAPSV 927

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 14 | 1.5 | 758 | 1 | US-08-256-188-2 |
| 2 | 14 | 1.5 | 758 | 1 | US-08-526-813-2 |
| 3 | 14 | 1.5 | 758 | 5 | PCT-US95-08554-2 |
| 4 | 14 | 1.5 | 1049 | 3 | US-08-772-270A-11 |
| 5 | 14 | 1.5 | 1053 | 4 | US-09-062-126-10 |
| 6 | 14 | 1.5 | 1244 | 5 | PCT-US93-10500-2 |
| 7 | 11 | 1.2 | 956 | 3 | US-08-772-270A-8 |
| 8 | 11 | 1.2 | 956 | 4 | US-09-062-126-8 |
| 9 | 11 | 1.2 | 1013 | 4 | US-09-062-126-3 |
| 10 | 11 | 1.2 | 1022 | 3 | US-08-772-270A-2 |
| 11 | 10 | 1.1 | 149 | 2 | US-08-387-942C-45 |
| 12 | 10 | 1.1 | 490 | 4 | US-09-252-149B-26 |
| 13 | 10 | 1.1 | 544 | 1 | US-08-387-156-10 |
| 14 | 10 | 1.1 | 544 | 2 | US-08-694-865-10 |
| 15 | 10 | 1.1 | 544 | 2 | US-08-878-748-10 |
| 16 | 10 | 1.1 | 544 | 3 | US-08-124-491-10 |
| 17 | 10 | 1.1 | 544 | 4 | US-08-383-912-10 |
| 18 | 10 | 1.1 | 699 | 2 | US-08-694-865-16 |
| 19 | 10 | 1.1 | 699 | 3 | US-09-124-491-16 |
| 20 | 10 | 1.1 | 699 | 4 | US-09-383-912-16 |
| 21 | 10 | 1.1 | 924 | 3 | US-08-619-812-8 |
| 22 | 10 | 1.1 | 926 | 1 | US-07-908-253-2 |
| 23 | 10 | 1.1 | 926 | 1 | US-08-455-970A-2 |
| 24 | 10 | 1.1 | 926 | 2 | US-08-387-156-6 |
| 25 | 10 | 1.1 | 926 | 2 | US-08-694-865-6 |
| 26 | 10 | 1.1 | 926 | 2 | US-08-878-748-6 |
| 27 | 10 | 1.1 | 926 | 2 | US-08-535-837-2 |
| 28 | 10 | 1.1 | 926 | 3 | US-09-124-491-6 |
| 29 | 10 | 1.1 | 926 | 4 | US-09-383-912-6 |
| 30 | 10 | 1.1 | 926 | 6 | 5476657-3 |
| 31 | 10 | 1.1 | 934 | 1 | US-08-215-805A-80 |
| 32 | 10 | 1.1 | 936 | 1 | US-08-455-970A-12 |
| 33 | 10 | 1.1 | 943 | 1 | US-08-455-970A-10 |
| 34 | 10 | 1.1 | 951 | 1 | US-08-455-970A-14 |
| 35 | 10 | 1.1 | 977 | 1 | US-08-387-156-8 |
| 36 | 10 | 1.1 | 977 | 2 | US-08-694-865-8 |
| 37 | 10 | 1.1 | 977 | 2 | US-08-878-748-8 |
| 38 | 10 | 1.1 | 977 | 3 | US-09-124-491-8 |
| 39 | 10 | 1.1 | 977 | 4 | US-09-383-912-8 |
| 40 | 10 | 1.1 | 1069 | 1 | US-07-777-715-9 |
| 41 | 10 | 1.1 | 1069 | 1 | US-08-170-126-4 |
| 42 | 10 | 1.1 | 1069 | 3 | US-08-954-418-4 |
| 43 | 10 | 1.1 | 1098 | 1 | US-07-777-715-7 |
| 44 | 10 | 1.1 | 1098 | 1 | US-08-170-126-2 |
| 45 | 10 | 1.1 | 1098 | 3 | US-08-954-418-2 |
| 46 | 10 | 1.1 | 1334 | 6 | 5476657-1 |
| 47 | 10 | 1.1 | 1403 | 1 | US-07-908-253-3 |
| 48 | 10 | 1.1 | 1403 | 2 | US-08-694-865-17 |
| 49 | 10 | 1.1 | 1403 | 2 | US-08-387-942C-3 |
| 50 | 10 | 1.1 | 1403 | 2 | US-08-535-837-3 |
| 51 | 10 | 1.1 | 1403 | 3 | US-09-124-491-17 |
| 52 | 10 | 1.1 | 1403 | 4 | US-09-383-912-17 |
| 53 | 8 | 0.9 | 128 | 4 | US-09-328-352-5276 |
| 54 | 8 | 0.9 | 142 | 2 | US-08-387-942C-41 |
| 55 | 8 | 0.9 | 151 | 2 | US-08-387-942C-46 |
| 56 | 8 | 0.9 | 151 | 2 | US-08-387-942C-50 |
| 57 | 8 | 0.9 | 153 | 2 | US-08-387-942C-47 |
| 58 | 8 | 0.9 | 153 | 2 | US-08-387-942C-51 |
| 59 | 8 | 0.9 | 355 | 4 | US-09-328-352-4560 |
| 60 | 8 | 0.9 | 499 | 4 | US-09-328-352-5089 |
| 61 | 8 | 0.9 | 553 | 2 | US-08-387-942C-2 |
| 62 | 8 | 0.9 | 805 | 4 | US-09-134-001C-4821 |
| 63 | 8 | 0.9 | 872 | 2 | US-08-387-942C-5 |
| 64 | 8 | 0.9 | 997 | 2 | US-08-387-942C-4 |
| 65 | 8 | 0.9 | 1181 | 4 | US-09-252-991A-18714 |
| 66 | 8 | 0.9 | 1705 | 4 | US-08-669-785-4 |
| 67 | 7 | 0.8 | 9 | 1 | US-08-215-805A-66 |
| 68 | 7 | 0.8 | 23 | 2 | US-08-295-643-11 |
| 69 | 7 | 0.8 | 23 | 3 | US-08-473-265-10 |
| 70 | 7 | 0.8 | 23 | 3 | US-08-284-747-10 |
| 71 | 7 | 0.8 | 25 | 1 | US-08-139-711-1 |
| 72 | 7 | 0.8 | 44 | 4 | US-08-469-260A-519 |
| 73 | 7 | 0.8 | 44 | 4 | US-08-488-446-519 |
| 74 | 7 | 0.8 | 44 | 4 | US-08-467-344A-519 |
| 75 | 7 | 0.8 | 52 | 3 | US-08-630-916A-78 |
| 76 | 7 | 0.8 | 103 | 4 | US-09-071-035-436 |
| 77 | 7 | 0.8 | 103 | 4 | US-09-252-991A-31182 |
| 78 | 7 | 0.8 | 103 | 4 | US-09-732-210-765 |
| 79 | 7 | 0.8 | 103 | 4 | US-09-732-210-766 |
| 80 | 7 | 0.8 | 104 | 4 | US-09-732-210-761 |
| 81 | 7 | 0.8 | 128 | 4 | US-09-071-035-434 |
| 82 | 7 | 0.8 | 151 | 4 | US-09-886-319A-58 |
| 83 | 7 | 0.8 | 151 | 2 | US-08-387-942C-49 |
| 84 | 7 | 0.8 | 153 | 4 | US-09-886-319A-57 |
| 85 | 7 | 0.8 | 210 | 3 | US-08-856-841-9 |
| 86 | 7 | 0.8 | 215 | 1 | US-08-044-621D-34 |
| 87 | 7 | 0.8 | 227 | 3 | US-09-252-991A-23558 |
| 88 | 7 | 0.8 | 233 | 4 | US-09-252-991A-27285 |
| 89 | 7 | 0.8 | 236 | 4 | US-09-252-991A-24211 |
| 90 | 7 | 0.8 | 242 | 4 | US-08-856-841-12 |
| 91 | 7 | 0.8 | 250 | 3 | US-08-532A-6895 |
| 92 | 7 | 0.8 | 263 | 4 | US-09-107-532A-6895 |
| 93 | 7 | 0.8 | 274 | 4 | US-09-266-965-125 |
| 94 | 7 | 0.8 | 308 | 4 | US-09-252-991A-23093 |
| 95 | 7 | 0.8 | 315 | 4 | US-09-134-001C-4521 |
| 96 | 7 | 0.8 | 337 | 4 | US-09-252-991A-26733 |
| 97 | 7 | 0.8 | 332 | 4 | US-09-107-532A-5477 |
| 98 | 7 | 0.8 | 338 | 4 | US-09-328-352-7151 |
| 99 | 7 | 0.8 | 373 | 4 | US-09-252-991A-31832 |
| 100 | 7 | 0.8 | 380 | 3 | US-09-150-133-9 |

Sequence 6, Appli
Sequence 6, Appli
Patent No. 5476657
Sequence 80, Appl
Sequence 12, Appl
Sequence 10, Appl
Sequence 14, Appl
Sequence 8, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 9, Appli
Sequence 4, Appli
Sequence 7, Appli
Sequence 2, Appli
Sequence 2, Appli
Patent No. 5476657
Sequence 3, Appli
Sequence 17, Appl
Sequence 3, Appli
Sequence 17, Appl
Sequence 17, Appl
Sequence 5276, Ap
Sequence 41, Appl
Sequence 46, Appl
Sequence 50, Appl
Sequence 47, Appl
Sequence 51, Appl
Sequence 4560, Ap
Sequence 5089, Ap
Sequence 2, Appli
Sequence 4821, Ap
Sequence 5, Appli
Sequence 4, Appli
Sequence 18714, A
Sequence 4, Appli
Sequence 66, Appl
Sequence 11, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 1, Appli
Sequence 519, App
Sequence 519, App
Sequence 519, App
Sequence 78, Appl
Sequence 436, Appl
Sequence 31182, A
Sequence 765, App
Sequence 766, App
Sequence 761, App
Sequence 434, App
Sequence 58, Appl
Sequence 49, Appl
Sequence 57, Appl
Sequence 9, Appli
Sequence 31, Appl
Sequence 15, Appl
Sequence 23558, A
Sequence 27285, A
Sequence 24211, A
Sequence 12, Appl
Sequence 125, App
Sequence 6895, App
Sequence 125, App
Sequence 23093, A
Sequence 4521, Ap
Sequence 26733, A
Sequence 5477, Ap
Sequence 7151, Ap
Sequence 31832, A
Sequence 9, Appli

| | | | | | | | | | | | | | |
|-----|---|-----|------|---|----------------------|-------------------|-----|---|-----|----|---|--------------------|--------------------|
| 101 | 7 | 0.8 | 380 | 3 | US-09-150-141-9 | Sequence 9, Appli | 174 | 6 | 0.6 | 12 | 2 | US-09-007-961-30 | Sequence 30, Appl |
| 102 | 7 | 0.8 | 380 | 3 | US-09-374-493-9 | Sequence 9, Appli | 175 | 6 | 0.6 | 12 | 3 | US-09-267-439-30 | Sequence 30, Appl |
| 103 | 7 | 0.8 | 380 | 3 | US-09-374-824-9 | Sequence 9, Appli | 176 | 6 | 0.6 | 12 | 3 | US-09-073-138-30 | Sequence 30, Appl |
| 104 | 7 | 0.8 | 380 | 3 | US-09-374-492-9 | Sequence 9, Appli | 177 | 6 | 0.6 | 13 | 1 | US-08-057-167-16 | Sequence 16, Appl |
| 105 | 7 | 0.8 | 387 | 3 | US-08-856-841-17 | Sequence 17, Appl | 178 | 6 | 0.6 | 13 | 5 | PCT-US93-05413-16 | Sequence 16, Appl |
| 106 | 7 | 0.8 | 389 | 4 | US-09-055-765-13 | Sequence 13, Appl | 179 | 6 | 0.6 | 15 | 2 | US-08-432-871C-81 | Sequence 81, Appl |
| 107 | 7 | 0.8 | 392 | 4 | US-09-107-532A-7217 | Sequence 7217, Ap | 180 | 6 | 0.6 | 15 | 4 | US-09-270-956-81 | Sequence 81, Appl |
| 108 | 7 | 0.8 | 424 | 4 | US-09-134-001C-5009 | Sequence 5009, Ap | 181 | 6 | 0.6 | 15 | 4 | US-09-517-866-6 | Sequence 6, Appli |
| 109 | 7 | 0.8 | 437 | 4 | US-09-107-532A-6709 | Sequence 6709, Ap | 182 | 6 | 0.6 | 15 | 6 | 5304631-8 | Patent No. 5304631 |
| 110 | 7 | 0.8 | 438 | 4 | US-09-252-991A-20127 | Sequence 20127, A | 183 | 6 | 0.6 | 15 | 6 | 5304631-16 | Patent No. 5304631 |
| 111 | 7 | 0.8 | 441 | 4 | US-09-107-532A-4173 | Sequence 4173, Ap | 184 | 6 | 0.6 | 16 | 3 | US-09-413-452-14 | Sequence 14, Appl |
| 112 | 7 | 0.8 | 446 | 4 | US-09-252-991A-18924 | Sequence 18924, A | 185 | 6 | 0.6 | 16 | 3 | US-09-413-068-14 | Sequence 14, Appl |
| 113 | 7 | 0.8 | 473 | 4 | US-09-252-991A-29636 | Sequence 29636, A | 186 | 6 | 0.6 | 17 | 4 | US-09-570-921-142 | Sequence 142, App |
| 114 | 7 | 0.8 | 477 | 3 | US-08-843-982B-8 | Sequence 8, Appli | 187 | 6 | 0.6 | 18 | 1 | US-08-086-335C-3 | Sequence 3, Appli |
| 115 | 7 | 0.8 | 478 | 3 | US-09-272-114-1 | Sequence 1, Appli | 188 | 6 | 0.6 | 18 | 3 | US-09-007-905-52 | Sequence 52, Appl |
| 116 | 7 | 0.8 | 478 | 3 | US-09-249-338-2 | Sequence 2, Appli | 189 | 6 | 0.6 | 18 | 4 | US-09-232-074-52 | Sequence 52, Appl |
| 117 | 7 | 0.8 | 480 | 4 | US-09-252-991A-26186 | Sequence 26186, A | 190 | 6 | 0.6 | 20 | 1 | US-07-777-715-4 | Sequence 4, Appli |
| 118 | 7 | 0.8 | 486 | 3 | US-08-856-841-19 | Sequence 19, Appl | 191 | 6 | 0.6 | 20 | 1 | US-07-908-253-5 | Sequence 5, Appli |
| 119 | 7 | 0.8 | 500 | 3 | US-08-856-841-16 | Sequence 16, Appl | 192 | 6 | 0.6 | 20 | 1 | US-08-086-335C-6 | Sequence 6, Appli |
| 120 | 7 | 0.8 | 500 | 4 | US-08-252-991A-25053 | Sequence 25053, A | 193 | 6 | 0.6 | 20 | 1 | US-08-170-126-6 | Sequence 6, Appli |
| 121 | 7 | 0.8 | 501 | 3 | US-08-856-841-18 | Sequence 18, Appl | 194 | 6 | 0.6 | 20 | 2 | US-08-535-837-5 | Sequence 5, Appli |
| 122 | 7 | 0.8 | 514 | 3 | US-09-413-814-82 | Sequence 82, Appl | 195 | 6 | 0.6 | 20 | 3 | US-08-954-418-6 | Sequence 6, Appli |
| 123 | 7 | 0.8 | 514 | 3 | US-09-413-814-94 | Sequence 94, Appl | 196 | 6 | 0.6 | 20 | 3 | US-08-825-852-51 | Sequence 51, Appl |
| 124 | 7 | 0.8 | 521 | 2 | US-08-721-684C-2 | Sequence 2, Appli | 197 | 6 | 0.6 | 20 | 3 | US-09-007-905-59 | Sequence 59, Appl |
| 125 | 7 | 0.8 | 521 | 2 | US-09-005-970-2 | Sequence 2, Appli | 198 | 6 | 0.6 | 20 | 3 | US-09-007-905-64 | Sequence 64, Appl |
| 126 | 7 | 0.8 | 521 | 3 | US-09-407-715-2 | Sequence 2, Appli | 199 | 6 | 0.6 | 20 | 3 | US-09-052-888-52 | Sequence 52, Appl |
| 127 | 7 | 0.8 | 523 | 4 | US-09-325-932A-183 | Sequence 183, App | 200 | 6 | 0.6 | 20 | 4 | US-09-232-074-59 | Sequence 59, Appl |
| 128 | 7 | 0.8 | 534 | 4 | US-09-252-991A-27300 | Sequence 27300, A | 201 | 6 | 0.6 | 20 | 4 | US-09-232-074-64 | Sequence 64, Appl |
| 129 | 7 | 0.8 | 557 | 3 | US-08-927-219-139 | Sequence 139, App | 202 | 6 | 0.6 | 20 | 6 | 5476657-6 | Patent No. 5476657 |
| 130 | 7 | 0.8 | 609 | 3 | US-08-927-219-129 | Sequence 129, App | 203 | 6 | 0.6 | 21 | 1 | US-08-086-335C-4 | Sequence 4, Appli |
| 131 | 7 | 0.8 | 679 | 4 | US-09-252-991A-28887 | Sequence 28887, A | 204 | 6 | 0.6 | 22 | 6 | 5489517-6 | Patent No. 5489517 |
| 132 | 7 | 0.8 | 707 | 4 | US-08-252-991A-24045 | Sequence 24045, A | 205 | 6 | 0.6 | 23 | 1 | US-08-141-324-23 | Sequence 23, Appl |
| 133 | 7 | 0.8 | 711 | 3 | US-08-949-588-2 | Sequence 2, Appli | 206 | 6 | 0.6 | 23 | 1 | US-08-541-902-23 | Sequence 23, Appl |
| 134 | 7 | 0.8 | 720 | 3 | US-09-286-284-25 | Sequence 25, Appl | 207 | 6 | 0.6 | 23 | 3 | US-08-448-194-16 | Sequence 16, Appl |
| 135 | 7 | 0.8 | 725 | 4 | US-09-252-991A-27825 | Sequence 27825, A | 208 | 6 | 0.6 | 23 | 4 | US-08-867-921-16 | Sequence 16, Appl |
| 136 | 7 | 0.8 | 748 | 4 | US-09-252-991A-31491 | Sequence 31491, A | 209 | 6 | 0.6 | 23 | 4 | US-09-570-921-65 | Sequence 65, Appl |
| 137 | 7 | 0.8 | 754 | 3 | US-09-296-284-4 | Sequence 4, Appli | 210 | 6 | 0.6 | 24 | 1 | US-08-133-011-9 | Sequence 9, Appli |
| 138 | 7 | 0.8 | 783 | 4 | US-08-252-991A-19698 | Sequence 19698, A | 211 | 6 | 0.6 | 24 | 1 | US-08-322-730A-9 | Sequence 9, Appli |
| 139 | 7 | 0.8 | 906 | 4 | US-09-328-352-6037 | Sequence 6037, Ap | 212 | 6 | 0.6 | 24 | 1 | US-08-387-874-9 | Sequence 9, Appli |
| 140 | 7 | 0.8 | 1021 | 4 | US-09-252-991A-27405 | Sequence 27405, A | 213 | 6 | 0.6 | 24 | 2 | US-08-468-790-10 | Sequence 10, Appl |
| 141 | 7 | 0.8 | 1048 | 4 | US-09-252-991A-29368 | Sequence 29368, A | 214 | 6 | 0.6 | 24 | 2 | US-08-383-619-9 | Sequence 9, Appli |
| 142 | 7 | 0.8 | 1075 | 2 | US-08-993-228-19 | Sequence 19, Appl | 215 | 6 | 0.6 | 24 | 3 | US-08-907-739-9 | Sequence 9, Appli |
| 143 | 7 | 0.8 | 1083 | 1 | US-08-895-601-5 | Sequence 5, Appli | 216 | 6 | 0.6 | 24 | 4 | US-09-729-597-9 | Sequence 9, Appli |
| 144 | 7 | 0.8 | 1138 | 1 | US-07-973-320-4 | Sequence 4, Appli | 217 | 6 | 0.6 | 24 | 4 | US-08-495-209-5 | Sequence 5, Appli |
| 145 | 7 | 0.8 | 1138 | 1 | US-07-973-320-4 | Sequence 4, Appli | 218 | 6 | 0.6 | 24 | 5 | PCT-US93-08364-9 | Sequence 9, Appli |
| 146 | 7 | 0.8 | 1242 | 4 | US-09-107-532A-5241 | Sequence 5241, Ap | 219 | 6 | 0.6 | 24 | 5 | PCT-US96-10905-5 | Sequence 5, Appli |
| 147 | 7 | 0.8 | 1287 | 5 | PCT-US95-02219-2 | Sequence 2, Appli | 220 | 6 | 0.6 | 25 | 1 | US-08-475-989-9 | Sequence 9, Appli |
| 148 | 7 | 0.8 | 1287 | 5 | PCT-US95-02219A-2 | Sequence 2, Appli | 221 | 6 | 0.6 | 25 | 2 | US-08-475-985-9 | Sequence 9, Appli |
| 149 | 7 | 0.8 | 1287 | 5 | US-08-470-260-3 | Sequence 3, Appli | 222 | 6 | 0.6 | 25 | 2 | US-08-472-172-21 | Sequence 21, Appl |
| 150 | 7 | 0.8 | 1296 | 3 | US-08-470-260-3 | Sequence 3, Appli | 223 | 6 | 0.6 | 25 | 3 | US-08-256-839-9 | Sequence 9, Appli |
| 151 | 7 | 0.8 | 1296 | 3 | US-08-471-491-3 | Sequence 3, Appli | 224 | 6 | 0.6 | 26 | 1 | US-07-971-096-12 | Sequence 12, Appl |
| 152 | 7 | 0.8 | 1296 | 3 | US-08-466-662-3 | Sequence 3, Appli | 225 | 6 | 0.6 | 26 | 1 | US-08-175-096-12 | Sequence 12, Appl |
| 153 | 7 | 0.8 | 1380 | 4 | US-09-328-352-8132 | Sequence 8132, Ap | 226 | 6 | 0.6 | 26 | 4 | US-09-461-325-383 | Sequence 383, App |
| 154 | 7 | 0.8 | 3135 | 1 | US-08-323-170B-2 | Sequence 2, Appli | 227 | 6 | 0.6 | 28 | 3 | US-09-413-452-15 | Sequence 15, Appl |
| 155 | 7 | 0.8 | 3135 | 4 | US-08-954-441-2 | Sequence 2, Appli | 228 | 6 | 0.6 | 28 | 3 | US-09-413-068-15 | Sequence 15, Appl |
| 156 | 6 | 0.6 | 8 | 3 | US-09-082-279B-1438 | Sequence 1438, Ap | 229 | 6 | 0.6 | 30 | 3 | US-08-448-489-7 | Sequence 7, Appli |
| 157 | 6 | 0.6 | 8 | 4 | US-09-834-784-1438 | Sequence 1438, Ap | 230 | 6 | 0.6 | 32 | 6 | 5464756-12 | Patent No. 5464756 |
| 158 | 6 | 0.6 | 9 | 1 | US-08-215-805A-3 | Sequence 3, Appli | 231 | 6 | 0.6 | 33 | 3 | US-08-810-712-25 | Sequence 25, Appl |
| 159 | 6 | 0.6 | 9 | 1 | US-08-215-805A-6 | Sequence 6, Appli | 232 | 6 | 0.6 | 33 | 4 | US-09-181-941-14 | Sequence 14, Appl |
| 160 | 6 | 0.6 | 9 | 1 | US-08-215-805A-7 | Sequence 7, Appli | 233 | 6 | 0.6 | 34 | 2 | US-08-248-839C-165 | Sequence 165, App |
| 161 | 6 | 0.6 | 9 | 1 | US-08-215-805A-9 | Sequence 9, Appli | 234 | 6 | 0.6 | 38 | 3 | US-09-007-905-48 | Sequence 48, Appl |
| 162 | 6 | 0.6 | 9 | 1 | US-08-215-805A-22 | Sequence 22, Appl | 235 | 6 | 0.6 | 38 | 4 | US-09-232-074-48 | Sequence 48, Appl |
| 163 | 6 | 0.6 | 9 | 1 | US-08-215-805A-24 | Sequence 24, Appl | 236 | 6 | 0.6 | 40 | 2 | US-08-637-759B-484 | Sequence 484, App |
| 164 | 6 | 0.6 | 9 | 1 | US-08-215-805A-27 | Sequence 27, Appl | 237 | 6 | 0.6 | 40 | 3 | US-08-871-355A-484 | Sequence 484, App |
| 165 | 6 | 0.6 | 9 | 1 | US-08-215-805A-32 | Sequence 32, Appl | 238 | 6 | 0.6 | 40 | 4 | US-09-201-945-484 | Sequence 484, App |
| 166 | 6 | 0.6 | 9 | 1 | US-08-215-805A-41 | Sequence 41, Appl | 239 | 6 | 0.6 | 43 | 3 | US-09-007-905-57 | Sequence 57, Appl |
| 167 | 6 | 0.6 | 9 | 1 | US-08-215-805A-62 | Sequence 62, Appl | 240 | 6 | 0.6 | 43 | 4 | US-09-232-074-57 | Sequence 57, Appl |
| 168 | 6 | 0.6 | 9 | 1 | US-08-215-805A-70 | Sequence 70, Appl | 241 | 6 | 0.6 | 45 | 3 | US-09-100-600A-87 | Sequence 87, Appl |
| 169 | 6 | 0.6 | 9 | 1 | US-08-215-805A-75 | Sequence 75, Appl | 242 | 6 | 0.6 | 45 | 3 | US-09-100-600A-90 | Sequence 90, Appl |
| 170 | 6 | 0.6 | 10 | 4 | PCT-US91-03056-19 | Sequence 19, Appl | 243 | 6 | 0.6 | 46 | 3 | US-08-857-076-83 | Sequence 83, Appl |
| 171 | 6 | 0.6 | 12 | 2 | US-09-517-866-9 | Sequence 9, Appli | 244 | 6 | 0.6 | 48 | 1 | US-08-062-472B-46 | Sequence 46, Appl |
| 172 | 6 | 0.6 | 12 | 2 | US-08-417-174-30 | Sequence 30, Appl | 245 | 6 | 0.6 | 48 | 4 | US-09-205-258-358 | Sequence 358, App |
| 173 | 6 | 0.6 | 12 | 2 | US-08-231-565A-30 | Sequence 30, Appl | 246 | 6 | 0.6 | 50 | 4 | US-09-447-125B-19 | Sequence 19, Appl |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|---|----------------------|--------------------|-----|---|-----|-----|---|----------------------|-------------------|
| 247 | 6 | 0.6 | 50 | 5 | PCT-US93-05640-33 | Sequence 33, Appl | 320 | 6 | 0.6 | 106 | 2 | US-08-246-361A-25 | Sequence 25, Appl |
| 248 | 6 | 0.6 | 51 | 1 | US-08-346-849-14 | Sequence 14, Appl | 321 | 6 | 0.6 | 106 | 3 | US-08-463-772-25 | Sequence 25, Appl |
| 249 | 6 | 0.6 | 51 | 2 | US-08-293-284A-14 | Sequence 14, Appl | 322 | 6 | 0.6 | 106 | 5 | PCT-US93-05000-25 | Sequence 25, Appl |
| 250 | 6 | 0.6 | 51 | 4 | US-09-187-789-34 | Sequence 34, Appl | 323 | 6 | 0.6 | 107 | 4 | US-09-134-001C-4474 | Sequence 4474, Ap |
| 251 | 6 | 0.6 | 51 | 4 | US-09-139-600-29 | Sequence 29, Appl | 324 | 6 | 0.6 | 107 | 4 | US-09-186-276B-63 | Sequence 63, Appl |
| 252 | 6 | 0.6 | 51 | 4 | US-08-998-600-14 | Sequence 14, Appl | 325 | 6 | 0.6 | 107 | 4 | US-08-842-445-63 | Sequence 63, Appl |
| 253 | 6 | 0.6 | 52 | 1 | US-08-247-475-39 | Sequence 39, Appl | 326 | 6 | 0.6 | 107 | 4 | US-09-186-188B-63 | Sequence 63, Appl |
| 254 | 6 | 0.6 | 52 | 1 | US-08-479-650-39 | Sequence 39, Appl | 327 | 6 | 0.6 | 111 | 1 | US-08-456-265A-111 | Sequence 111, App |
| 255 | 6 | 0.6 | 52 | 1 | US-08-191-866D-61 | Sequence 61, Appl | 328 | 6 | 0.6 | 111 | 2 | US-08-971-217-111 | Sequence 111, App |
| 256 | 6 | 0.6 | 52 | 1 | US-08-674-169-39 | Sequence 39, Appl | 329 | 6 | 0.6 | 111 | 3 | US-09-350-600-111 | Sequence 111, App |
| 257 | 6 | 0.6 | 52 | 2 | US-08-185-949B-61 | Sequence 61, Appl | 330 | 6 | 0.6 | 111 | 4 | US-08-469-260A-72 | Sequence 72, Appl |
| 258 | 6 | 0.6 | 58 | 4 | US-08-963-851-22 | Sequence 22, Appl | 331 | 6 | 0.6 | 111 | 4 | US-08-488-446-72 | Sequence 72, Appl |
| 259 | 6 | 0.6 | 59 | 3 | US-08-812-586-7 | Sequence 7, Appli | 332 | 6 | 0.6 | 111 | 4 | US-08-467-344A-72 | Sequence 22828, A |
| 260 | 6 | 0.6 | 59 | 4 | US-09-535-832A-7 | Sequence 7, Appli | 333 | 6 | 0.6 | 112 | 4 | US-09-252-991A-23828 | Sequence 17, Appl |
| 261 | 6 | 0.6 | 60 | 2 | US-08-417-210A-113 | Sequence 113, App | 334 | 6 | 0.6 | 115 | 1 | US-08-209-747-17 | Sequence 17, Appl |
| 262 | 6 | 0.6 | 60 | 3 | US-08-812-586-6 | Sequence 6, Appli | 335 | 6 | 0.6 | 115 | 1 | US-08-458-298-17 | Sequence 17, Appl |
| 263 | 6 | 0.6 | 60 | 4 | US-09-535-832A-6 | Sequence 6, Appli | 336 | 6 | 0.6 | 116 | 1 | US-08-209-747-18 | Sequence 18, Appl |
| 264 | 6 | 0.6 | 61 | 4 | US-09-134-001C-4272 | Sequence 4272, Ap | 337 | 6 | 0.6 | 116 | 1 | US-08-458-298-18 | Sequence 18, Appl |
| 265 | 6 | 0.6 | 63 | 3 | US-08-817-787-17 | Sequence 17, Appl | 338 | 6 | 0.6 | 116 | 4 | US-09-252-991A-23793 | Sequence 23793, A |
| 266 | 6 | 0.6 | 63 | 6 | 5198542-11 | Patent No. 5198542 | 339 | 6 | 0.6 | 116 | 4 | US-09-732-210-661 | Sequence 661, App |
| 267 | 6 | 0.6 | 64 | 4 | US-09-205-258-570 | Sequence 570, App | 340 | 6 | 0.6 | 118 | 4 | US-09-732-210-158 | Sequence 158, App |
| 268 | 6 | 0.6 | 65 | 3 | US-09-007-905-36 | Sequence 36, Appl | 341 | 6 | 0.6 | 119 | 4 | US-09-328-352-5881 | Sequence 5881, Ap |
| 269 | 6 | 0.6 | 65 | 4 | US-09-232-074-36 | Sequence 36, Appl | 342 | 6 | 0.6 | 120 | 2 | US-08-637-759B-269 | Sequence 269, App |
| 270 | 6 | 0.6 | 66 | 3 | US-09-007-905-40 | Sequence 40, Appl | 343 | 6 | 0.6 | 120 | 3 | US-08-871-355A-269 | Sequence 269, App |
| 271 | 6 | 0.6 | 66 | 4 | US-09-232-074-40 | Sequence 40, Appl | 344 | 6 | 0.6 | 120 | 4 | US-09-201-945-269 | Sequence 269, App |
| 272 | 6 | 0.6 | 69 | 3 | US-09-007-905-30 | Sequence 30, Appl | 345 | 6 | 0.6 | 122 | 4 | US-09-462-606-65 | Sequence 65, Appl |
| 273 | 6 | 0.6 | 69 | 4 | US-09-232-074-30 | Sequence 30, Appl | 346 | 6 | 0.6 | 123 | 3 | US-08-840-316-3 | Sequence 3, Appli |
| 274 | 6 | 0.6 | 75 | 4 | US-09-252-991A-28889 | Sequence 28889, A | 347 | 6 | 0.6 | 123 | 3 | US-08-478-507-9 | Sequence 9, Appli |
| 275 | 6 | 0.6 | 75 | 6 | 5252466-3 | Patent No. 5252466 | 348 | 6 | 0.6 | 123 | 3 | US-08-809-523-3 | Sequence 3, Appli |
| 276 | 6 | 0.6 | 77 | 4 | US-08-936-165A-511 | Sequence 511, App | 349 | 6 | 0.6 | 123 | 3 | US-09-128-275A-9 | Sequence 9, Appli |
| 277 | 6 | 0.6 | 81 | 4 | US-09-198-452A-1167 | Sequence 1167, Ap | 350 | 6 | 0.6 | 123 | 3 | US-08-471-971-3 | Sequence 3, Appli |
| 278 | 6 | 0.6 | 82 | 4 | US-09-107-532A-4732 | Sequence 4732, Ap | 351 | 6 | 0.6 | 123 | 4 | US-09-553-427-9 | Sequence 9, Appli |
| 279 | 6 | 0.6 | 83 | 2 | US-08-858-207A-272 | Sequence 272, App | 352 | 6 | 0.6 | 123 | 4 | US-09-462-606-13 | Sequence 13, Appl |
| 280 | 6 | 0.6 | 83 | 2 | US-08-858-767-33 | Sequence 33, Appl | 353 | 6 | 0.6 | 123 | 4 | US-09-462-606-58 | Sequence 58, Appl |
| 281 | 6 | 0.6 | 83 | 2 | US-08-858-767-34 | Sequence 34, Appl | 354 | 6 | 0.6 | 123 | 4 | US-09-462-606-59 | Sequence 59, Appl |
| 282 | 6 | 0.6 | 83 | 2 | US-08-863-028-33 | Sequence 33, Appl | 355 | 6 | 0.6 | 123 | 4 | US-09-462-606-60 | Sequence 60, Appl |
| 283 | 6 | 0.6 | 83 | 2 | US-08-863-028-34 | Sequence 34, Appl | 356 | 6 | 0.6 | 123 | 4 | US-09-462-606-61 | Sequence 61, Appl |
| 284 | 6 | 0.6 | 87 | 4 | US-09-191-468-45 | Sequence 45, Appl | 357 | 6 | 0.6 | 123 | 4 | US-09-462-606-62 | Sequence 62, Appl |
| 285 | 6 | 0.6 | 87 | 4 | US-09-191-468-47 | Sequence 47, Appl | 358 | 6 | 0.6 | 123 | 4 | US-09-462-606-63 | Sequence 63, Appl |
| 286 | 6 | 0.6 | 87 | 4 | US-09-191-468-49 | Sequence 49, Appl | 359 | 6 | 0.6 | 123 | 4 | US-09-462-606-64 | Sequence 64, Appl |
| 287 | 6 | 0.6 | 87 | 4 | US-09-191-468-51 | Sequence 51, Appl | 360 | 6 | 0.6 | 123 | 4 | US-08-469-260A-61 | Sequence 61, Appl |
| 288 | 6 | 0.6 | 87 | 4 | US-09-191-468-53 | Sequence 53, Appl | 361 | 6 | 0.6 | 123 | 4 | US-09-402-776-3 | Sequence 3, Appli |
| 289 | 6 | 0.6 | 90 | 4 | US-09-328-352-6160 | Sequence 6160, Ap | 362 | 6 | 0.6 | 123 | 4 | US-09-172-699-4 | Sequence 4, Appli |
| 290 | 6 | 0.6 | 91 | 3 | US-08-905-223-409 | Sequence 409, App | 363 | 6 | 0.6 | 123 | 4 | US-08-488-446-61 | Sequence 61, Appl |
| 291 | 6 | 0.6 | 94 | 1 | US-08-286-451B-20 | Sequence 20, Appl | 364 | 6 | 0.6 | 123 | 4 | US-08-467-344A-61 | Sequence 61, Appl |
| 292 | 6 | 0.6 | 94 | 2 | US-08-748-725-20 | Sequence 20, Appl | 365 | 6 | 0.6 | 123 | 5 | PCT-US93-08849A-3 | Sequence 3, Appli |
| 293 | 6 | 0.6 | 94 | 2 | US-09-047-125-32 | Sequence 32, Appl | 366 | 6 | 0.6 | 123 | 5 | PCT-US93-08849-3 | Sequence 3, Appli |
| 294 | 6 | 0.6 | 94 | 3 | US-07-736-335B-32 | Sequence 32, Appl | 367 | 6 | 0.6 | 124 | 1 | US-08-240-049B-19 | Sequence 19, Appl |
| 295 | 6 | 0.6 | 94 | 4 | US-09-461-325-381 | Sequence 381, App | 368 | 6 | 0.6 | 124 | 1 | US-08-240-049B-20 | Sequence 20, Appl |
| 296 | 6 | 0.6 | 95 | 1 | US-08-456-265A-107 | Sequence 107, App | 369 | 6 | 0.6 | 124 | 3 | US-08-542-634-21 | Sequence 21, Appl |
| 297 | 6 | 0.6 | 95 | 1 | US-08-456-265A-109 | Sequence 109, App | 370 | 6 | 0.6 | 124 | 3 | US-08-542-634-22 | Sequence 22, Appl |
| 298 | 6 | 0.6 | 95 | 1 | US-08-456-265A-110 | Sequence 110, App | 371 | 6 | 0.6 | 124 | 3 | US-08-477-292-21 | Sequence 21, Appl |
| 299 | 6 | 0.6 | 95 | 2 | US-08-971-217-107 | Sequence 107, App | 372 | 6 | 0.6 | 124 | 3 | US-08-477-292-22 | Sequence 22, Appl |
| 300 | 6 | 0.6 | 95 | 2 | US-08-971-217-109 | Sequence 109, App | 373 | 6 | 0.6 | 124 | 4 | US-09-252-991A-31009 | Sequence 31009, A |
| 301 | 6 | 0.6 | 95 | 2 | US-08-971-217-110 | Sequence 110, App | 374 | 6 | 0.6 | 124 | 5 | PCT-US95-13703-21 | Sequence 21, Appl |
| 302 | 6 | 0.6 | 95 | 3 | US-09-350-600-107 | Sequence 107, App | 375 | 6 | 0.6 | 124 | 5 | PCT-US95-13703-22 | Sequence 22, Appl |
| 303 | 6 | 0.6 | 95 | 3 | US-09-350-600-109 | Sequence 109, App | 376 | 6 | 0.6 | 127 | 3 | US-08-806-121B-3 | Sequence 3, Appli |
| 304 | 6 | 0.6 | 95 | 3 | US-09-350-600-110 | Sequence 110, App | 377 | 6 | 0.6 | 127 | 3 | US-08-812-586-47 | Sequence 47, Appl |
| 305 | 6 | 0.6 | 98 | 4 | US-09-107-532A-7167 | Sequence 7167, Ap | 378 | 6 | 0.6 | 127 | 4 | US-09-443-061-3 | Sequence 3, Appli |
| 306 | 6 | 0.6 | 99 | 2 | US-08-537-400-19 | Sequence 19, Appl | 379 | 6 | 0.6 | 127 | 4 | US-09-535-832A-44 | Sequence 44, Appl |
| 307 | 6 | 0.6 | 99 | 4 | US-09-370-838-95 | Sequence 95, Appl | 380 | 6 | 0.6 | 127 | 4 | US-09-732-210-1685 | Sequence 1685, Ap |
| 308 | 6 | 0.6 | 100 | 2 | US-08-464-517-8 | Sequence 8, Appli | 381 | 6 | 0.6 | 127 | 4 | US-09-107-532A-6784 | Sequence 6784, Ap |
| 309 | 6 | 0.6 | 100 | 2 | US-08-246-361A-8 | Sequence 8, Appli | 382 | 6 | 0.6 | 129 | 3 | US-09-476-482-8 | Sequence 8, Appli |
| 310 | 6 | 0.6 | 100 | 3 | US-08-463-772-8 | Sequence 8, Appli | 383 | 6 | 0.6 | 130 | 2 | US-08-771-602D-44 | Sequence 44, Appl |
| 311 | 6 | 0.6 | 100 | 5 | PCT-US93-05000-8 | Patent No. 5164490 | 384 | 6 | 0.6 | 132 | 1 | US-07-820-154A-2 | Sequence 2, Appli |
| 312 | 6 | 0.6 | 101 | 6 | 5164490-5 | Sequence 5, Appl | 385 | 6 | 0.6 | 132 | 1 | US-07-820-154A-6 | Sequence 6, Appli |
| 313 | 6 | 0.6 | 103 | 4 | US-09-732-210-763 | Sequence 763, App | 386 | 6 | 0.6 | 132 | 2 | US-08-097-554A-2 | Sequence 2, Appli |
| 314 | 6 | 0.6 | 103 | 4 | US-09-732-210-1231 | Sequence 1231, Ap | 387 | 6 | 0.6 | 132 | 2 | US-08-097-554A-6 | Sequence 6, Appli |
| 315 | 6 | 0.6 | 104 | 3 | US-08-857-076-112 | Sequence 112, App | 388 | 6 | 0.6 | 132 | 3 | US-08-480-640A-2 | Sequence 2, Appli |
| 316 | 6 | 0.6 | 105 | 3 | US-08-867-381A-4 | Sequence 4, Appli | 389 | 6 | 0.6 | 132 | 3 | US-08-480-640A-6 | Sequence 6, Appli |
| 317 | 6 | 0.6 | 105 | 4 | US-09-521-144-4 | Sequence 781, App | 390 | 6 | 0.6 | 132 | 3 | US-08-295-802-2 | Sequence 2, Appli |
| 318 | 6 | 0.6 | 105 | 4 | US-09-732-210-781 | Sequence 781, App | 391 | 6 | 0.6 | 132 | 3 | US-08-295-802-6 | Sequence 6, Appli |
| 319 | 6 | 0.6 | 106 | 2 | US-08-464-517-25 | Sequence 25, Appl | 392 | 6 | 0.6 | 132 | 3 | US-08-686-968C-102 | Sequence 102, App |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|---|----------------------|--------------------|-----|---|-----|-----|---|----------------------|--------------------|
| 393 | 6 | 0.6 | 132 | 3 | US-08-488-237A-2 | Sequence 2, Appli | 466 | 6 | 0.6 | 157 | 1 | US-08-450-595-2 | Sequence 2, Appli |
| 394 | 6 | 0.6 | 132 | 3 | US-08-488-237A-6 | Sequence 6, Appli | 467 | 6 | 0.6 | 157 | 3 | US-08-818-562-2 | Sequence 2, Appli |
| 395 | 6 | 0.6 | 132 | 4 | US-08-375-992A-2 | Sequence 2, Appli | 468 | 6 | 0.6 | 157 | 4 | US-09-134-001C-2991 | Sequence 2991, Ap |
| 396 | 6 | 0.6 | 132 | 4 | US-08-375-992A-6 | Sequence 6, Appli | 469 | 6 | 0.6 | 157 | 4 | US-09-628-445-2 | Sequence 2, Appli |
| 397 | 6 | 0.6 | 132 | 4 | US-08-472-679H-2 | Sequence 2, Appli | 470 | 6 | 0.6 | 158 | 2 | US-08-888-497-22 | Sequence 22, Appl |
| 398 | 6 | 0.6 | 132 | 4 | US-08-472-679H-6 | Sequence 6, Appli | 471 | 6 | 0.6 | 158 | 2 | US-09-362-230-22 | Sequence 22, Appl |
| 399 | 6 | 0.6 | 132 | 5 | PCT-US93-00324-2 | Sequence 2, Appli | 472 | 6 | 0.6 | 158 | 4 | US-09-252-991A-17285 | Sequence 17285, A |
| 400 | 6 | 0.6 | 132 | 5 | PCT-US93-00324-6 | Sequence 6, Appli | 473 | 6 | 0.6 | 158 | 4 | US-09-252-991A-28490 | Sequence 28490, A |
| 401 | 6 | 0.6 | 133 | 1 | US-07-800-366-1 | Sequence 1, Appli | 474 | 6 | 0.6 | 158 | 5 | PCT-US94-07926-22 | Sequence 22, Appl |
| 402 | 6 | 0.6 | 133 | 1 | US-08-354-456A-5 | Sequence 5, Appli | 475 | 6 | 0.6 | 159 | 4 | US-09-732-210-727 | Sequence 727, Ap |
| 403 | 6 | 0.6 | 133 | 1 | US-08-354-456A-6 | Sequence 6, Appli | 476 | 6 | 0.6 | 160 | 4 | US-09-134-001C-5659 | Sequence 5659, Ap |
| 404 | 6 | 0.6 | 133 | 1 | US-08-225-224-3 | Sequence 3, Appli | 477 | 6 | 0.6 | 162 | 2 | US-08-716-317-10 | Sequence 10, Appl |
| 405 | 6 | 0.6 | 133 | 1 | US-08-318-133-89 | Sequence 89, Appl | 478 | 6 | 0.6 | 163 | 2 | US-08-716-317-15 | Sequence 15, Appl |
| 406 | 6 | 0.6 | 133 | 1 | US-08-284-393B-1 | Sequence 1, Appli | 479 | 6 | 0.6 | 163 | 2 | US-08-846-762-75 | Sequence 75, Appl |
| 407 | 6 | 0.6 | 133 | 1 | US-08-284-393B-2 | Sequence 2, Appli | 480 | 6 | 0.6 | 163 | 3 | US-08-990-791-11 | Sequence 11, Appl |
| 408 | 6 | 0.6 | 133 | 1 | US-08-284-393B-3 | Sequence 3, Appli | 481 | 6 | 0.6 | 163 | 4 | US-09-724-623-107 | Sequence 107, App |
| 409 | 6 | 0.6 | 133 | 1 | US-08-734-471-1 | Sequence 1, Appli | 482 | 6 | 0.6 | 163 | 4 | US-09-372-591-11 | Sequence 11, Appl |
| 410 | 6 | 0.6 | 133 | 3 | US-08-722-258-3 | Sequence 3, Appli | 483 | 6 | 0.6 | 164 | 1 | US-08-318-193-60 | Sequence 60, Appl |
| 411 | 6 | 0.6 | 133 | 3 | US-08-905-223-112 | Sequence 412, App | 484 | 6 | 0.6 | 164 | 4 | US-09-252-991A-16845 | Sequence 16845, A |
| 412 | 6 | 0.6 | 133 | 3 | US-08-817-787-13 | Sequence 13, Appl | 485 | 6 | 0.6 | 166 | 4 | US-09-499-148-1 | Sequence 1, Appli |
| 413 | 6 | 0.6 | 133 | 4 | US-08-570-921-54 | Sequence 54, Appl | 486 | 6 | 0.6 | 166 | 4 | US-09-499-148-4 | Sequence 4, Appli |
| 414 | 6 | 0.6 | 133 | 4 | US-09-252-991A-29319 | Sequence 29319, A | 487 | 6 | 0.6 | 167 | 1 | US-08-246-427A-2 | Sequence 2, Appli |
| 415 | 6 | 0.6 | 133 | 4 | US-09-538-873-2 | Sequence 2, Appli | 488 | 6 | 0.6 | 167 | 2 | US-08-766-620-2 | Sequence 2, Appli |
| 416 | 6 | 0.6 | 133 | 5 | PCT-US95-04468-3 | Sequence 3, Appli | 489 | 6 | 0.6 | 167 | 3 | US-08-959-004-7 | Sequence 7, Appli |
| 417 | 6 | 0.6 | 133 | 5 | PCT-US95-08950-1 | Sequence 1, Appli | 490 | 6 | 0.6 | 167 | 5 | PCT-US95-06094-2 | Sequence 2, Appli |
| 418 | 6 | 0.6 | 133 | 5 | PCT-US95-08950-2 | Sequence 2, Appli | 491 | 6 | 0.6 | 168 | 2 | US-08-353-476-73 | Sequence 73, Appl |
| 419 | 6 | 0.6 | 133 | 5 | PCT-US95-08950-3 | Sequence 3, Appli | 492 | 6 | 0.6 | 168 | 4 | US-09-252-991A-27144 | Sequence 27144, A |
| 420 | 6 | 0.6 | 133 | 5 | PCT-US95-08950-3 | Patent No. 5210029 | 493 | 6 | 0.6 | 168 | 4 | US-09-252-991A-32471 | Sequence 32471, A |
| 421 | 6 | 0.6 | 133 | 6 | 5256769-1 | Patent No. 5256769 | 494 | 6 | 0.6 | 168 | 4 | US-09-328-352-7984 | Sequence 7984, Ap |
| 422 | 6 | 0.6 | 133 | 6 | 5464939-2 | Patent No. 5464939 | 495 | 6 | 0.6 | 171 | 4 | US-09-134-001C-3256 | Sequence 3256, Ap |
| 423 | 6 | 0.6 | 134 | 6 | 5496324-55 | Patent No. 5496324 | 496 | 6 | 0.6 | 171 | 4 | US-09-328-352-4243 | Sequence 4243, Ap |
| 424 | 6 | 0.6 | 135 | 4 | US-08-858-207A-489 | Sequence 489, App | 497 | 6 | 0.6 | 172 | 3 | US-08-149-101A-21 | Sequence 21, Appl |
| 425 | 6 | 0.6 | 136 | 2 | US-08-580-545B-8 | Sequence 8, Appli | 498 | 6 | 0.6 | 172 | 5 | PCT-US94-12873-21 | Sequence 21, Appl |
| 426 | 6 | 0.6 | 136 | 3 | US-09-262-653A-8 | Sequence 8, Appli | 499 | 6 | 0.6 | 173 | 1 | US-08-193-977-4 | Sequence 4, Appli |
| 427 | 6 | 0.6 | 138 | 4 | US-09-328-352-7972 | Sequence 7972, Ap | 500 | 6 | 0.6 | 173 | 4 | US-09-107-532A-5804 | Sequence 5804, Ap |
| 428 | 6 | 0.6 | 141 | 1 | US-07-695-564-5 | Sequence 5, Appli | 501 | 6 | 0.6 | 174 | 3 | US-08-149-101A-20 | Sequence 20, Appl |
| 429 | 6 | 0.6 | 141 | 1 | US-08-241-387-5 | Sequence 5, Appli | 502 | 6 | 0.6 | 174 | 4 | US-09-252-991A-21312 | Sequence 21312, A |
| 430 | 6 | 0.6 | 141 | 2 | US-08-690-011A-43 | Sequence 43, Appl | 503 | 6 | 0.6 | 174 | 5 | PCT-US94-12873-20 | Sequence 20, Appl |
| 431 | 6 | 0.6 | 141 | 4 | US-09-299-495F-43 | Sequence 43, Appl | 504 | 6 | 0.6 | 176 | 1 | US-07-924-054-8 | Sequence 8, Appli |
| 432 | 6 | 0.6 | 141 | 4 | US-09-732-210-758 | Sequence 758, App | 505 | 6 | 0.6 | 176 | 1 | US-08-076-011-2 | Sequence 2, Appli |
| 433 | 6 | 0.6 | 143 | 4 | US-09-252-991A-26817 | Sequence 26817, A | 506 | 6 | 0.6 | 176 | 1 | US-08-062-472B-20 | Sequence 20, Appl |
| 434 | 6 | 0.6 | 145 | 2 | US-08-686-599A-20 | Sequence 20, Appl | 507 | 6 | 0.6 | 176 | 3 | US-09-232-446B-4 | Sequence 4, Appli |
| 435 | 6 | 0.6 | 145 | 4 | US-09-134-001C-5243 | Sequence 5243, Ap | 508 | 6 | 0.6 | 176 | 6 | 5198542-4 | Patent No. 5198542 |
| 436 | 6 | 0.6 | 146 | 4 | US-09-252-991A-22315 | Sequence 22315, A | 509 | 6 | 0.6 | 178 | 3 | US-09-232-446B-5 | Sequence 5, Appli |
| 437 | 6 | 0.6 | 146 | 4 | US-09-252-991A-25506 | Sequence 25506, A | 510 | 6 | 0.6 | 179 | 1 | US-08-193-977-6 | Sequence 6, Appli |
| 438 | 6 | 0.6 | 147 | 4 | US-09-252-991A-17088 | Sequence 17088, A | 511 | 6 | 0.6 | 179 | 3 | US-08-867-381A-52 | Sequence 52, Appl |
| 439 | 6 | 0.6 | 148 | 1 | US-08-207-904-15 | Sequence 15, Appl | 512 | 6 | 0.6 | 179 | 4 | US-09-521-144-52 | Sequence 52, Appl |
| 440 | 6 | 0.6 | 149 | 1 | US-07-695-564-7 | Sequence 7, Appli | 513 | 6 | 0.6 | 181 | 1 | US-08-209-182C-4 | Sequence 4, Appli |
| 441 | 6 | 0.6 | 149 | 1 | US-08-241-387-7 | Sequence 7, Appli | 514 | 6 | 0.6 | 181 | 3 | US-08-961-083-114 | Sequence 114, App |
| 442 | 6 | 0.6 | 149 | 4 | US-08-252-991A-30582 | Sequence 30582, A | 515 | 6 | 0.6 | 181 | 4 | US-09-536-784-114 | Sequence 114, App |
| 443 | 6 | 0.6 | 150 | 2 | US-08-460-694-3 | Sequence 3, Appli | 516 | 6 | 0.6 | 181 | 5 | PCT-US92-05614-4 | Sequence 4, Appli |
| 444 | 6 | 0.6 | 150 | 2 | US-08-387-942C-44 | Sequence 44, Appl | 517 | 6 | 0.6 | 182 | 3 | US-08-149-101A-25 | Sequence 25, Appl |
| 445 | 6 | 0.6 | 150 | 2 | US-08-387-942C-52 | Sequence 52, Appl | 518 | 6 | 0.6 | 182 | 5 | PCT-US94-12873-25 | Sequence 25, Appl |
| 446 | 6 | 0.6 | 150 | 3 | US-08-460-744-3 | Sequence 3, Appli | 519 | 6 | 0.6 | 183 | 1 | US-08-009-973-1 | Sequence 1, Appli |
| 447 | 6 | 0.6 | 150 | 3 | US-07-667-711B-3 | Sequence 3, Appli | 520 | 6 | 0.6 | 184 | 1 | US-08-567-047-2 | Sequence 2, Appli |
| 448 | 6 | 0.6 | 150 | 3 | US-08-211-542A-14 | Sequence 14, Appl | 521 | 6 | 0.6 | 184 | 2 | US-08-693-182-2 | Sequence 2, Appli |
| 449 | 6 | 0.6 | 151 | 2 | US-08-387-942C-42 | Sequence 42, Appl | 522 | 6 | 0.6 | 184 | 2 | US-08-567-048-2 | Sequence 2, Appli |
| 450 | 6 | 0.6 | 152 | 4 | US-09-107-532A-3744 | Sequence 3744, Ap | 523 | 6 | 0.6 | 184 | 2 | US-09-008-482-2 | Sequence 2, Appli |
| 451 | 6 | 0.6 | 153 | 2 | US-08-387-942C-10 | Sequence 10, Appl | 524 | 6 | 0.6 | 184 | 2 | US-08-945-529-8 | Sequence 8, Appli |
| 452 | 6 | 0.6 | 153 | 2 | US-08-387-942C-43 | Sequence 43, Appl | 525 | 6 | 0.6 | 184 | 2 | US-08-945-529-9 | Sequence 9, Appli |
| 453 | 6 | 0.6 | 153 | 3 | US-09-012-366-3 | Sequence 3, Appli | 526 | 6 | 0.6 | 184 | 2 | US-08-945-529-10 | Sequence 10, Appl |
| 454 | 6 | 0.6 | 153 | 3 | US-08-759-628-8 | Sequence 8, Appli | 527 | 6 | 0.6 | 184 | 2 | US-08-945-529-11 | Sequence 11, Appl |
| 455 | 6 | 0.6 | 153 | 4 | US-09-522-217-111 | Sequence 111, App | 528 | 6 | 0.6 | 184 | 2 | US-08-945-529-12 | Sequence 12, Appl |
| 456 | 6 | 0.6 | 153 | 4 | US-09-328-352-4423 | Sequence 4423, Ap | 529 | 6 | 0.6 | 184 | 3 | US-08-149-101A-17 | Sequence 17, Appl |
| 457 | 6 | 0.6 | 153 | 6 | 5314995-7 | Patent No. 5314995 | 530 | 6 | 0.6 | 184 | 3 | US-08-149-101A-22 | Sequence 22, Appl |
| 458 | 6 | 0.6 | 154 | 4 | US-09-252-991A-29366 | Sequence 29366, A | 531 | 6 | 0.6 | 184 | 5 | US-09-252-991A-26628 | Sequence 26628, A |
| 459 | 6 | 0.6 | 154 | 4 | US-09-107-532A-4031 | Sequence 4031, Ap | 532 | 6 | 0.6 | 184 | 5 | PCT-US94-12873-17 | Sequence 17, Appl |
| 460 | 6 | 0.6 | 155 | 1 | US-08-150-203A-8 | Sequence 8, Appli | 533 | 6 | 0.6 | 184 | 6 | PCT-US94-12873-22 | Sequence 22, Appl |
| 461 | 6 | 0.6 | 155 | 1 | US-08-454-730-8 | Sequence 8, Appli | 534 | 6 | 0.6 | 184 | 6 | 5186931-1 | Patent No. 5186931 |
| 462 | 6 | 0.6 | 155 | 4 | US-08-949-788-8 | Sequence 8, Appli | 535 | 6 | 0.6 | 185 | 1 | US-07-632-070B-1 | Sequence 1, Appli |
| 463 | 6 | 0.6 | 155 | 4 | US-09-252-991A-26113 | Sequence 26113, A | 536 | 6 | 0.6 | 185 | 1 | US-07-918-181A-2 | Sequence 2, Appli |
| 464 | 6 | 0.6 | 156 | 4 | US-09-134-001C-5252 | Sequence 5252, Ap | 537 | 6 | 0.6 | 185 | 1 | US-07-918-181A-4 | Sequence 4, Appli |
| 465 | 6 | 0.6 | 157 | 1 | US-08-450-065-2 | Sequence 2, Appli | 538 | 6 | 0.6 | 185 | 1 | US-07-918-181A-6 | Sequence 6, Appli |

| | | | | | | | | | | | |
|-----|---|-----|-----|---|--------------------|--------------------|-----|-----|---|----------------------|--------------------|
| 539 | 6 | 0.6 | 185 | 1 | US-07-918-181A-8 | Sequence 8, Appli | 612 | 193 | 4 | US-09-328-352-6034 | Sequence 6034, Ap |
| 540 | 6 | 0.6 | 185 | 1 | US-08-231-575-2 | Sequence 2, Appli | 613 | 194 | 1 | US-08-063-552-8 | Sequence 8, Appli |
| 541 | 6 | 0.6 | 185 | 1 | US-08-231-575-4 | Sequence 4, Appli | 614 | 194 | 5 | PCT-US93-05704-8 | Sequence 8, Appli |
| 542 | 6 | 0.6 | 185 | 1 | US-08-231-575-6 | Sequence 6, Appli | 615 | 195 | 4 | US-09-252-991A-18622 | Sequence 18622, A |
| 543 | 6 | 0.6 | 185 | 1 | US-08-231-575-8 | Sequence 8, Appli | 616 | 196 | 4 | US-09-252-991A-19503 | Sequence 19503, A |
| 544 | 6 | 0.6 | 185 | 1 | US-08-246-427A-5 | Sequence 5, Appli | 617 | 196 | 4 | US-09-252-991A-26430 | Sequence 26430, A |
| 545 | 6 | 0.6 | 185 | 2 | US-08-716-317-7 | Sequence 7, Appli | 618 | 197 | 4 | US-09-125-619-25 | Sequence 25, Appl |
| 546 | 6 | 0.6 | 185 | 2 | US-08-766-620-5 | Sequence 5, Appli | 619 | 197 | 4 | US-09-252-991A-21791 | Sequence 21791, A |
| 547 | 6 | 0.6 | 185 | 3 | US-08-705-875A-5 | Sequence 5, Appli | 620 | 199 | 4 | US-09-252-991A-31679 | Sequence 31679, A |
| 548 | 6 | 0.6 | 185 | 3 | US-09-122-443-16 | Sequence 16, Appl | 621 | 199 | 4 | US-09-107-532A-6773 | Sequence 6773, Ap |
| 549 | 6 | 0.6 | 185 | 3 | US-09-220-731-22 | Sequence 22, Appl | 622 | 200 | 1 | US-07-855-412B-1 | Sequence 1, Appli |
| 550 | 6 | 0.6 | 185 | 4 | US-09-558-089-16 | Sequence 16, Appl | 623 | 200 | 1 | US-08-209-182C-8 | Sequence 8, Appli |
| 551 | 6 | 0.6 | 185 | 4 | US-09-558-087-16 | Sequence 16, Appl | 624 | 200 | 2 | US-08-308-887A-1 | Sequence 1, Appli |
| 552 | 6 | 0.6 | 185 | 4 | US-09-242-999-5 | Sequence 5, Appli | 625 | 200 | 3 | US-08-881-094-1 | Sequence 1, Appli |
| 553 | 6 | 0.6 | 185 | 4 | US-09-198-452A-106 | Sequence 106, App | 626 | 200 | 4 | US-09-252-991A-18724 | Sequence 18724, A |
| 554 | 6 | 0.6 | 185 | 5 | PCT-US92-05612-2 | Sequence 2, Appli | 627 | 200 | 4 | US-09-252-991A-27855 | Sequence 27855, A |
| 555 | 6 | 0.6 | 185 | 5 | PCT-US93-06928-2 | Sequence 2, Appli | 628 | 201 | 4 | US-09-252-991A-21340 | Sequence 21340, A |
| 556 | 6 | 0.6 | 185 | 5 | PCT-US93-06928-4 | Sequence 4, Appli | 629 | 201 | 5 | PCT-US92-05612-8 | Sequence 8, Appli |
| 557 | 6 | 0.6 | 185 | 5 | PCT-US93-06928-6 | Sequence 6, Appli | 630 | 203 | 6 | 5168051-12 | Patent No. 5168051 |
| 558 | 6 | 0.6 | 185 | 5 | PCT-US93-06928-8 | Sequence 8, Appli | 631 | 204 | 3 | US-09-230-637-25 | Sequence 25, Appl |
| 559 | 6 | 0.6 | 185 | 6 | 5186931-2 | Patent No. 5186931 | 632 | 204 | 4 | US-09-230-371A-26 | Sequence 26, Appl |
| 560 | 6 | 0.6 | 186 | 1 | US-07-632-070B-2 | Sequence 2, Appli | 633 | 205 | 4 | US-09-107-532A-6070 | Sequence 6070, Ap |
| 561 | 6 | 0.6 | 186 | 1 | US-07-745-382-20 | Sequence 20, Appl | 634 | 206 | 4 | US-09-252-991A-18746 | Sequence 18746, A |
| 562 | 6 | 0.6 | 186 | 1 | US-07-921-848-20 | Sequence 20, Appl | 635 | 207 | 4 | US-09-199-637A-211 | Sequence 211, App |
| 563 | 6 | 0.6 | 186 | 1 | US-08-209-182C-2 | Sequence 2, Appli | 636 | 208 | 3 | US-08-630-172-15 | Sequence 15, Appl |
| 564 | 6 | 0.6 | 186 | 1 | US-08-209-182C-6 | Sequence 6, Appli | 637 | 208 | 3 | US-09-375-419-15 | Sequence 15, Appl |
| 565 | 6 | 0.6 | 186 | 1 | US-08-185-301A-20 | Sequence 20, Appl | 638 | 208 | 4 | US-09-460-384-36 | Sequence 36, Appl |
| 566 | 6 | 0.6 | 186 | 3 | US-08-469-318-163 | Sequence 163, App | 639 | 208 | 4 | US-09-710-092-8 | Sequence 8, Appli |
| 567 | 6 | 0.6 | 186 | 3 | US-08-468-609A-163 | Sequence 163, App | 640 | 209 | 4 | US-09-252-991A-18156 | Sequence 18156, A |
| 568 | 6 | 0.6 | 186 | 3 | US-08-810-436-20 | Sequence 20, Appl | 641 | 209 | 4 | US-09-328-352-8077 | Sequence 8077, A |
| 569 | 6 | 0.6 | 186 | 4 | US-08-446-827A-163 | Sequence 163, App | 642 | 210 | 4 | US-08-897-537A-2 | Sequence 2, Appli |
| 570 | 6 | 0.6 | 186 | 4 | US-08-762-827A-163 | Sequence 163, App | 643 | 212 | 1 | US-08-792-019B-9 | Sequence 9, Appli |
| 571 | 6 | 0.6 | 186 | 5 | PCT-US92-05612-6 | Sequence 6, Appli | 644 | 212 | 3 | US-08-988-819-9 | Sequence 9, Appli |
| 572 | 6 | 0.6 | 186 | 5 | PCT-US94-14179-20 | Sequence 20, Appl | 645 | 212 | 3 | US-09-016-534-9 | Sequence 9, Appli |
| 573 | 6 | 0.6 | 186 | 5 | PCT-US95-01185-163 | Sequence 163, App | 646 | 212 | 3 | US-08-097-869-7 | Sequence 7, Appli |
| 574 | 6 | 0.6 | 187 | 1 | US-07-632-070B-3 | Sequence 3, Appli | 647 | 212 | 3 | US-08-795-473B-6 | Sequence 6, Appli |
| 575 | 6 | 0.6 | 188 | 3 | US-09-122-443-12 | Sequence 12, Appl | 648 | 212 | 3 | US-09-230-637-45 | Sequence 45, Appl |
| 576 | 6 | 0.6 | 188 | 4 | US-09-558-089-12 | Sequence 12, Appl | 649 | 212 | 3 | US-09-155-941-1 | Sequence 1, Appli |
| 577 | 6 | 0.6 | 188 | 4 | US-09-558-087-12 | Sequence 12, Appl | 650 | 212 | 4 | US-09-230-371A-27 | Sequence 27, Appl |
| 578 | 6 | 0.6 | 189 | 1 | US-08-152-485-4 | Sequence 4, Appli | 651 | 212 | 4 | US-09-439-856-6 | Sequence 6, Appli |
| 579 | 6 | 0.6 | 189 | 1 | US-08-463-089-4 | Sequence 4, Appli | 652 | 212 | 4 | US-09-487-792-14 | Sequence 14, Appl |
| 580 | 6 | 0.6 | 189 | 1 | US-08-461-360A-4 | Sequence 4, Appli | 653 | 212 | 4 | US-09-908-594-14 | Sequence 14, Appl |
| 581 | 6 | 0.6 | 189 | 1 | US-08-461-359-4 | Sequence 4, Appli | 654 | 212 | 6 | 5510472-2 | Patent No. 5510472 |
| 582 | 6 | 0.6 | 189 | 4 | US-08-671-548C-52 | Sequence 52, Appl | 655 | 213 | 4 | US-09-252-991A-22585 | Sequence 22585, A |
| 583 | 6 | 0.6 | 189 | 5 | PCT-US94-12904-4 | Sequence 4, Appli | 656 | 215 | 4 | US-09-134-001C-4736 | Sequence 4736, Ap |
| 584 | 6 | 0.6 | 191 | 3 | US-08-764-563-1 | Sequence 1, Appli | 657 | 215 | 4 | US-09-252-991A-27377 | Sequence 27377, A |
| 585 | 6 | 0.6 | 191 | 3 | US-08-720-625-2 | Sequence 2, Appli | 658 | 215 | 4 | US-09-252-991A-31700 | Sequence 31700, A |
| 586 | 6 | 0.6 | 191 | 4 | US-09-183-861-82 | Sequence 82, Appl | 659 | 219 | 1 | US-08-152-485-2 | Sequence 2, Appli |
| 587 | 6 | 0.6 | 191 | 4 | US-09-022-765-82 | Sequence 82, Appl | 660 | 219 | 1 | US-08-463-089-2 | Sequence 2, Appli |
| 588 | 6 | 0.6 | 191 | 4 | US-09-551-974A-82 | Sequence 82, Appl | 661 | 219 | 1 | US-08-461-360A-2 | Sequence 2, Appli |
| 589 | 6 | 0.6 | 192 | 1 | US-08-086-428B-81 | Sequence 81, Appl | 662 | 219 | 1 | US-08-461-359-2 | Sequence 2, Appli |
| 590 | 6 | 0.6 | 192 | 1 | US-08-086-428B-82 | Sequence 82, Appl | 663 | 219 | 5 | PCT-US94-12904-2 | Sequence 2, Appli |
| 591 | 6 | 0.6 | 192 | 1 | US-08-086-428B-83 | Sequence 83, Appl | 664 | 221 | 4 | US-09-328-352-5680 | Sequence 5680, Ap |
| 592 | 6 | 0.6 | 192 | 1 | US-08-086-428B-84 | Sequence 84, Appl | 665 | 221 | 4 | US-09-107-532A-4477 | Sequence 4477, Ap |
| 593 | 6 | 0.6 | 192 | 2 | US-08-468-570-81 | Sequence 81, Appl | 666 | 221 | 4 | US-09-107-532A-5877 | Sequence 5877, Ap |
| 594 | 6 | 0.6 | 192 | 2 | US-08-468-570-82 | Sequence 82, Appl | 667 | 222 | 4 | US-09-107-532A-4165 | Sequence 4165, Ap |
| 595 | 6 | 0.6 | 192 | 2 | US-08-468-570-83 | Sequence 83, Appl | 668 | 223 | 2 | US-08-528-612-3 | Sequence 3, Appli |
| 596 | 6 | 0.6 | 192 | 2 | US-08-468-570-84 | Sequence 84, Appl | 669 | 224 | 4 | US-09-082-920-2 | Sequence 2, Appli |
| 597 | 6 | 0.6 | 192 | 2 | US-08-230-665A-81 | Sequence 81, Appl | 670 | 225 | 1 | US-08-290-979A-8 | Sequence 8, Appli |
| 598 | 6 | 0.6 | 192 | 2 | US-08-230-665A-82 | Sequence 82, Appl | 671 | 226 | 4 | US-09-134-001C-5494 | Sequence 5494, Ap |
| 599 | 6 | 0.6 | 192 | 2 | US-08-230-665A-83 | Sequence 83, Appl | 672 | 226 | 4 | US-09-489-847-215 | Sequence 215, App |
| 600 | 6 | 0.6 | 192 | 2 | US-08-230-665A-84 | Sequence 84, Appl | 673 | 227 | 4 | US-09-252-991A-30173 | Sequence 30173, A |
| 601 | 6 | 0.6 | 192 | 3 | US-09-370-807-10 | Sequence 10, Appl | 674 | 227 | 4 | US-09-198-452A-1245 | Sequence 1245, Ap |
| 602 | 6 | 0.6 | 192 | 4 | US-08-466-601A-81 | Sequence 81, Appl | 675 | 229 | 4 | US-08-936-165A-523 | Sequence 523, App |
| 603 | 6 | 0.6 | 192 | 4 | US-08-466-601A-82 | Sequence 82, Appl | 676 | 229 | 4 | US-09-107-532A-4842 | Sequence 4842, Ap |
| 604 | 6 | 0.6 | 192 | 4 | US-08-466-601A-83 | Sequence 83, Appl | 677 | 233 | 4 | US-09-252-991A-25399 | Sequence 25399, A |
| 605 | 6 | 0.6 | 192 | 4 | US-08-466-601A-84 | Sequence 84, Appl | 678 | 234 | 4 | US-09-252-991A-21695 | Sequence 21695, A |
| 606 | 6 | 0.6 | 192 | 4 | US-08-466-601A-85 | Sequence 85, Appl | 679 | 235 | 1 | US-08-015-985-13 | Sequence 13, Appl |
| 607 | 6 | 0.6 | 192 | 5 | PCT-US93-10398-81 | Sequence 81, Appl | 680 | 235 | 2 | US-08-378-939-12 | Sequence 12, Appl |
| 608 | 6 | 0.6 | 192 | 5 | PCT-US95-10398-82 | Sequence 82, Appl | 681 | 235 | 3 | US-09-198-452A-752 | Sequence 752, App |
| 609 | 6 | 0.6 | 192 | 5 | PCT-US95-10398-83 | Sequence 83, Appl | 682 | 236 | 3 | US-08-937-271-11 | Sequence 11, Appl |
| 610 | 6 | 0.6 | 192 | 5 | PCT-US95-10398-84 | Sequence 84, Appl | 683 | 237 | 2 | US-08-578-709-11 | Sequence 11, Appl |
| 611 | 6 | 0.6 | 193 | 1 | US-08-248-466B-14 | Sequence 14, Appl | 684 | 237 | 4 | US-09-328-352-8181 | Sequence 8181, Ap |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|---|----------------------|--------------------|-----|---|-----|-----|---|----------------------|--------------------|
| 685 | 6 | 0.6 | 239 | 4 | US-09-372-422A-42 | Sequence 42, Appl | 758 | 6 | 0.6 | 273 | 2 | US-08-353-476-112 | Sequence 112, Appl |
| 686 | 6 | 0.6 | 239 | 4 | US-09-252-991A-19895 | Sequence 19895, A | 759 | 6 | 0.6 | 273 | 3 | US-08-928-213B-14 | Sequence 14, Appl |
| 687 | 6 | 0.6 | 239 | 4 | US-09-252-991A-25446 | Sequence 25446, A | 760 | 6 | 0.6 | 273 | 4 | US-09-252-991A-22948 | Sequence 22948, A |
| 688 | 6 | 0.6 | 240 | 4 | US-09-489-847-143 | Sequence 143, Appl | 761 | 6 | 0.6 | 275 | 4 | US-09-847-224A-26 | Sequence 26, Appl |
| 689 | 6 | 0.6 | 240 | 4 | US-09-252-991A-26041 | Sequence 26041, A | 762 | 6 | 0.6 | 277 | 1 | US-08-024-868-2 | Sequence 2, Appl |
| 690 | 6 | 0.6 | 241 | 2 | US-08-353-476-116 | Sequence 116, Appl | 763 | 6 | 0.6 | 277 | 1 | US-08-890-542A-2 | Sequence 2, Appl |
| 691 | 6 | 0.6 | 242 | 3 | US-09-067-800-2 | Sequence 2, Appl | 764 | 6 | 0.6 | 277 | 2 | US-08-242-097-2 | Sequence 2, Appl |
| 692 | 6 | 0.6 | 242 | 3 | US-09-105-852-2 | Sequence 2, Appl | 765 | 6 | 0.6 | 277 | 2 | US-08-567-375-16 | Sequence 16, Appl |
| 693 | 6 | 0.6 | 242 | 3 | US-09-349-677-2 | Sequence 2, Appl | 766 | 6 | 0.6 | 277 | 3 | US-08-591-605-2 | Sequence 2, Appl |
| 694 | 6 | 0.6 | 242 | 3 | US-09-708-584-2 | Sequence 2, Appl | 767 | 6 | 0.6 | 277 | 3 | US-08-964-308-6 | Sequence 6, Appl |
| 695 | 6 | 0.6 | 242 | 4 | US-09-252-991A-19136 | Sequence 19136, A | 768 | 6 | 0.6 | 277 | 3 | US-08-964-308-10 | Sequence 10, Appl |
| 696 | 6 | 0.6 | 246 | 4 | US-09-328-352-5613 | Sequence 5613, Ap | 769 | 6 | 0.6 | 277 | 3 | US-08-462-969B-4 | Sequence 4, Appl |
| 697 | 6 | 0.6 | 246 | 2 | US-08-851-974-3 | Sequence 3, Appl | 770 | 6 | 0.6 | 277 | 3 | US-08-964-313-6 | Sequence 6, Appl |
| 698 | 6 | 0.6 | 248 | 2 | US-09-213-390-3 | Sequence 3, Appl | 771 | 6 | 0.6 | 277 | 3 | US-08-964-313-10 | Sequence 10, Appl |
| 699 | 6 | 0.6 | 248 | 3 | US-08-750-145A-22 | Sequence 22, Appl | 772 | 6 | 0.6 | 277 | 3 | US-09-206-695-2 | Sequence 2, Appl |
| 700 | 6 | 0.6 | 248 | 3 | US-08-975-698A-26 | Sequence 26, Appl | 773 | 6 | 0.6 | 277 | 4 | US-09-000-179-1 | Sequence 1, Appl |
| 701 | 6 | 0.6 | 248 | 3 | US-09-417-090-26 | Sequence 26, Appl | 774 | 6 | 0.6 | 277 | 4 | US-09-069-138-6 | Sequence 6, Appl |
| 702 | 6 | 0.6 | 248 | 3 | US-09-727-578-26 | Sequence 26, Appl | 775 | 6 | 0.6 | 277 | 4 | US-09-069-138-10 | Sequence 10, Appl |
| 703 | 6 | 0.6 | 248 | 4 | US-09-252-991A-24788 | Sequence 24788, A | 776 | 6 | 0.6 | 277 | 4 | US-09-561-756-12 | Sequence 12, Appl |
| 704 | 6 | 0.6 | 251 | 1 | US-08-209-747-8 | Sequence 8, Appl | 777 | 6 | 0.6 | 277 | 4 | US-09-227-721-12 | Sequence 12, Appl |
| 705 | 6 | 0.6 | 251 | 1 | US-08-458-298-8 | Sequence 8, Appl | 778 | 6 | 0.6 | 277 | 4 | US-08-983-502-30 | Sequence 30, Appl |
| 706 | 6 | 0.6 | 251 | 3 | US-08-875-811-59 | Sequence 59, Appl | 779 | 6 | 0.6 | 277 | 4 | US-09-124-934A-4 | Sequence 4, Appl |
| 707 | 6 | 0.6 | 252 | 2 | US-08-685-992-33 | Sequence 33, Appl | 780 | 6 | 0.6 | 277 | 4 | US-08-724-378D-5 | Sequence 5, Appl |
| 708 | 6 | 0.6 | 252 | 2 | US-09-144-925-33 | Sequence 33, Appl | 781 | 6 | 0.6 | 277 | 4 | US-09-799-118-2 | Sequence 2, Appl |
| 709 | 6 | 0.6 | 252 | 3 | US-08-812-586-48 | Sequence 48, Appl | 782 | 6 | 0.6 | 277 | 4 | US-08-334-251D-4 | Sequence 4, Appl |
| 710 | 6 | 0.6 | 252 | 4 | US-09-489-847-214 | Sequence 214, Appl | 783 | 6 | 0.6 | 277 | 4 | US-09-198-452A-10 | Sequence 10, Appl |
| 711 | 6 | 0.6 | 252 | 4 | US-09-535-832A-45 | Sequence 45, Appl | 784 | 6 | 0.6 | 277 | 4 | US-09-516-747-30 | Sequence 30, Appl |
| 712 | 6 | 0.6 | 252 | 4 | US-09-252-991A-26557 | Sequence 26557, A | 785 | 6 | 0.6 | 277 | 5 | PCT-US96-10521-30 | Sequence 30, Appl |
| 713 | 6 | 0.6 | 253 | 3 | US-09-392-014-4 | Sequence 4, Appl | 786 | 6 | 0.6 | 277 | 5 | PCT-US96-11995-1 | Sequence 1, Appl |
| 714 | 6 | 0.6 | 253 | 4 | US-09-328-352-6440 | Sequence 6440, Ap | 787 | 6 | 0.6 | 278 | 1 | US-08-188-582-9 | Sequence 9, Appl |
| 715 | 6 | 0.6 | 254 | 3 | US-08-875-811-61 | Sequence 61, Appl | 788 | 6 | 0.6 | 278 | 1 | US-08-646-715-9 | Sequence 9, Appl |
| 716 | 6 | 0.6 | 254 | 4 | US-09-134-001C-4565 | Sequence 4565, Ap | 789 | 6 | 0.6 | 279 | 1 | US-08-312-387B-8 | Sequence 8, Appl |
| 717 | 6 | 0.6 | 254 | 4 | US-08-914-479A-4 | Sequence 4, Appl | 790 | 6 | 0.6 | 279 | 1 | US-08-683-426-8 | Sequence 8, Appl |
| 718 | 6 | 0.6 | 254 | 4 | US-09-252-991A-29819 | Sequence 29819, A | 791 | 6 | 0.6 | 279 | 2 | US-08-683-458-8 | Sequence 8, Appl |
| 719 | 6 | 0.6 | 255 | 4 | US-08-858-207A-465 | Sequence 465, App | 792 | 6 | 0.6 | 279 | 2 | US-08-878-360-8 | Sequence 8, Appl |
| 720 | 6 | 0.6 | 255 | 4 | US-09-107-532A-5335 | Sequence 5335, Ap | 793 | 6 | 0.6 | 279 | 3 | US-09-008-960-1 | Sequence 1, Appl |
| 721 | 6 | 0.6 | 256 | 2 | US-09-008-960-3 | Sequence 3, Appl | 794 | 6 | 0.6 | 279 | 3 | US-09-368-240-1 | Sequence 1, Appl |
| 722 | 6 | 0.6 | 256 | 3 | US-09-368-240-3 | Sequence 3, Appl | 795 | 6 | 0.6 | 279 | 4 | US-09-468-702-1 | Sequence 1, Appl |
| 723 | 6 | 0.6 | 256 | 4 | US-09-468-702-3 | Sequence 3, Appl | 796 | 6 | 0.6 | 279 | 4 | US-09-333-412-8 | Sequence 8, Appl |
| 724 | 6 | 0.6 | 257 | 1 | US-07-956-700B-105 | Sequence 105, Appl | 797 | 6 | 0.6 | 279 | 4 | US-09-328-352-4229 | Sequence 4229, Ap |
| 725 | 6 | 0.6 | 257 | 1 | US-08-476-537-105 | Sequence 105, Appl | 798 | 6 | 0.6 | 279 | 5 | PCT-US91-00899-8 | Sequence 8, Appl |
| 726 | 6 | 0.6 | 257 | 1 | US-08-485-607-105 | Sequence 105, Appl | 799 | 6 | 0.6 | 280 | 2 | US-08-872-961A-5 | Sequence 5, Appl |
| 727 | 6 | 0.6 | 257 | 2 | US-08-637-759B-92 | Sequence 92, Appl | 800 | 6 | 0.6 | 280 | 3 | US-09-231-258-5 | Sequence 5, Appl |
| 728 | 6 | 0.6 | 257 | 2 | US-08-475-879-105 | Sequence 105, Appl | 801 | 6 | 0.6 | 280 | 4 | US-09-194-146-10 | Sequence 10, Appl |
| 729 | 6 | 0.6 | 257 | 3 | US-08-871-355A-92 | Sequence 92, Appl | 802 | 6 | 0.6 | 280 | 4 | US-09-328-352-4975 | Sequence 4975, Ap |
| 730 | 6 | 0.6 | 257 | 4 | US-09-201-945-92 | Sequence 92, Appl | 803 | 6 | 0.6 | 282 | 2 | US-08-284-941-8 | Sequence 8, Appl |
| 731 | 6 | 0.6 | 257 | 4 | US-09-134-001C-3562 | Sequence 3562, Ap | 804 | 6 | 0.6 | 282 | 2 | US-08-447-642-8 | Sequence 8, Appl |
| 732 | 6 | 0.6 | 257 | 4 | US-09-433-043B-105 | Sequence 105, Appl | 805 | 6 | 0.6 | 282 | 3 | US-09-236-503-8 | Sequence 8, Appl |
| 733 | 6 | 0.6 | 257 | 4 | US-09-252-991A-26380 | Sequence 26380, A | 806 | 6 | 0.6 | 282 | 4 | US-09-198-452A-739 | Sequence 739, App |
| 734 | 6 | 0.6 | 257 | 4 | US-09-328-352-6621 | Sequence 6621, Ap | 807 | 6 | 0.6 | 282 | 4 | US-09-107-532A-5553 | Sequence 5553, Ap |
| 735 | 6 | 0.6 | 257 | 4 | US-09-107-532A-5440 | Sequence 5440, Ap | 808 | 6 | 0.6 | 282 | 4 | PCT-US93-02147A-8 | Sequence 8, Appl |
| 736 | 6 | 0.6 | 258 | 1 | US-07-721-761A-29 | Sequence 29, Appl | 809 | 6 | 0.6 | 283 | 3 | US-08-867-381A-3 | Sequence 3, Appl |
| 737 | 6 | 0.6 | 258 | 1 | US-07-978-687-29 | Sequence 29, Appl | 810 | 6 | 0.6 | 283 | 4 | US-09-521-144-3 | Sequence 3, Appl |
| 738 | 6 | 0.6 | 258 | 5 | PCT-US91-05801-29 | Sequence 29, Appl | 811 | 6 | 0.6 | 283 | 4 | US-09-134-001C-5346 | Sequence 5346, Ap |
| 739 | 6 | 0.6 | 260 | 4 | US-09-252-991A-24755 | Sequence 24755, A | 812 | 6 | 0.6 | 283 | 4 | US-09-134-001C-5491 | Sequence 5491, Ap |
| 740 | 6 | 0.6 | 261 | 4 | US-09-489-847-371 | Sequence 371, Appl | 813 | 6 | 0.6 | 283 | 4 | US-09-609-816-8 | Sequence 8, Appl |
| 741 | 6 | 0.6 | 263 | 4 | US-09-252-991A-27114 | Sequence 27114, A | 814 | 6 | 0.6 | 283 | 4 | US-09-609-816-9 | Sequence 9, Appl |
| 742 | 6 | 0.6 | 263 | 4 | US-09-252-991A-27644 | Sequence 27644, A | 815 | 6 | 0.6 | 284 | 4 | US-08-914-479A-6 | Sequence 6, Appl |
| 743 | 6 | 0.6 | 265 | 4 | US-09-996-243-223 | Sequence 223, App | 816 | 6 | 0.6 | 284 | 4 | US-09-561-756-35 | Sequence 35, Appl |
| 744 | 6 | 0.6 | 268 | 3 | US-09-353-585-6 | Sequence 6, Appl | 817 | 6 | 0.6 | 285 | 4 | US-09-227-721-35 | Sequence 35, Appl |
| 745 | 6 | 0.6 | 268 | 4 | US-09-252-991A-17449 | Sequence 17449, A | 818 | 6 | 0.6 | 285 | 4 | US-09-107-532A-6136 | Sequence 6136, Ap |
| 746 | 6 | 0.6 | 269 | 4 | US-09-328-352-6139 | Sequence 6139, Ap | 819 | 6 | 0.6 | 287 | 2 | US-08-424-641B-10 | Sequence 10, Appl |
| 747 | 6 | 0.6 | 270 | 4 | US-09-205-258-568 | Sequence 568, App | 820 | 6 | 0.6 | 287 | 2 | US-08-820-980-10 | Sequence 10, Appl |
| 748 | 6 | 0.6 | 270 | 4 | US-09-252-991A-30244 | Sequence 30244, A | 821 | 6 | 0.6 | 287 | 2 | US-08-826-439-10 | Sequence 10, Appl |
| 749 | 6 | 0.6 | 271 | 4 | US-09-252-991A-17292 | Sequence 17292, A | 822 | 6 | 0.6 | 287 | 4 | US-09-252-991A-21605 | Sequence 21605, A |
| 750 | 6 | 0.6 | 271 | 4 | US-09-252-991A-23450 | Sequence 23450, A | 823 | 6 | 0.6 | 287 | 4 | US-09-252-991A-23091 | Sequence 23091, A |
| 751 | 6 | 0.6 | 271 | 4 | US-09-252-991A-32772 | Sequence 32772, A | 824 | 6 | 0.6 | 287 | 4 | US-09-252-991A-32232 | Sequence 32232, A |
| 752 | 6 | 0.6 | 271 | 4 | US-09-107-532A-5157 | Sequence 5157, Ap | 825 | 6 | 0.6 | 288 | 2 | US-08-147-772-2 | Sequence 2, Appl |
| 753 | 6 | 0.6 | 272 | 1 | US-08-117-083-12 | Sequence 12, Appl | 826 | 6 | 0.6 | 288 | 2 | US-08-456-104-6 | Sequence 6, Appl |
| 754 | 6 | 0.6 | 272 | 2 | US-08-709-874A-2 | Sequence 2, Appl | 827 | 6 | 0.6 | 288 | 2 | US-08-101-624-23 | Sequence 23, Appl |
| 755 | 6 | 0.6 | 272 | 3 | US-09-104-382-2 | Sequence 2, Appl | 828 | 6 | 0.6 | 288 | 2 | US-08-751-767A-6 | Sequence 6, Appl |
| 756 | 6 | 0.6 | 272 | 4 | US-09-372-422A-26 | Sequence 26, Appl | 829 | 6 | 0.6 | 288 | 3 | US-08-153-262-2 | Sequence 2, Appl |
| 757 | 6 | 0.6 | 272 | 4 | US-09-328-352-7372 | Sequence 7372, Ap | 830 | 6 | 0.6 | 288 | 3 | US-08-479-744A-29 | Sequence 29, Appl |

977 Sequence 857, App
978 Patent No. 5210183
979 Sequence 5, Appl
980 Sequence 21150, A
981 Sequence 19, Appl
982 Sequence 19, Appl
983 Sequence 19, Appl
984 Sequence 5508, Ap
985 Sequence 18559, A
986 Sequence 2, Appl
987 Sequence 2, Appl
988 Sequence 24803, A
989 Sequence 18268, A
990 Sequence 29676, A
991 Sequence 37, App
992 Sequence 6481, Ap
993 Sequence 4325, Ap
994 Sequence 23, Appl
995 Sequence 27528, A
996 Sequence 9, Appl
997 Sequence 3899, Ap
998 Sequence 536, App
999 Sequence 4747, Ap
1000 Sequence 7825, Ap

ALIGNMENTS

RESULT 1
US-08-258-188-2
; Sequence 2, Application US/08258188
; Patent No. 5475098
; GENERAL INFORMATION:
; APPLICANT: HALL, Robert H.
; TITLE OF INVENTION: A NEW AND DISTINCTIVE DNA SEQUENCE OF E.
; TITLE OF INVENTION: coli O157:H7 AND ITS USE FOR THE RAPID, SENSITIVE AND
; TITLE OF INVENTION: SPECIFIC DETECTION OF O157:H7 AND OTHER ENTEROHEMORRHAGIC
; TITLE OF INVENTION: E. coli
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/258,188
; FILING DATE: 14-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 15280206, DHHSE135940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 758 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-258-188-2

Query Match 1.5%; Score 14; DB 1; Length 758;

Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 379 ISGILEASKQAMFE 392
Db 152 ISGILEASKQAMFE 165
|||||
RESULT 2
US-08-526-813-2
; Sequence 2, Application US/08526813
; Patent No. 5756293
; GENERAL INFORMATION:
; APPLICANT: Hall, Robert H.
; APPLICANT: Xu, Jian Guo
; TITLE OF INVENTION: A New and Distinctive DNA Sequence of E.
; TITLE OF INVENTION: coli O157:H7 and its Use for the Rapid, Sensitive and
; TITLE OF INVENTION: Specific Detection of O157:H7 and Other Enterohemorrhagic
; TITLE OF INVENTION: E. coli
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,813
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US 08/258,188
; FILING DATE: 14-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 15280-206-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 758 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-526-813-2
Query Match 1.5%; Score 14; DB 1; Length 758;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 379 ISGILEASKQAMFE 392
Db 152 ISGILEASKQAMFE 165
|||||
RESULT 3
PCT-US95-08554-2
; Sequence 2, Application PC/TUS9508554
; GENERAL INFORMATION:
; APPLICANT: Hall, Robert H.
; APPLICANT: Xu, Jian Guo
; TITLE OF INVENTION: A NEW AND DISTINCTIVE DNA SEQUENCE OF E.
; TITLE OF INVENTION: coli O157:H7 AND ITS USE FOR THE RAPID, SENSITIVE AND
; TITLE OF INVENTION: SPECIFIC DETECTION OF O157:H7 AND OTHER ENTEROHEMORRHAGIC

```
; TITLE OF INVENTION: E. coli
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08554
; FILING DATE: 14-JUN-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 15280206, DHHSE135940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 758 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-08554-2

Query Match 1.5%; Score 14; DB 5; Length 758;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 379 ISGILEASKQAMPE 392
Db 152 ISGILEASKQAMFE 165

RESULT 4
US-08-772-270A-11
; Sequence 11, Application US/08772270A
; Patent No. 6019984
; GENERAL INFORMATION:
; APPLICANT: MacInnes, Janet
; APPLICANT: Ricciatti, Paul
; APPLICANT: Mallard, Bonnie
; APPLICANT: Rosendal, Soren
; TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
; TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772-270A
; FILING DATE: December 23, 1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
```

```
; REFERENCE/DOCKET NUMBER: 6580-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1049 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Actinobacillus pleuropneumoniae
; US-08-772-270A-11

Query Match 1.5%; Score 14; DB 3; Length 1049;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 KVAAGFELSNOVIG 254
Db 275 KVAAGFELSNOVIG 288

RESULT 5
US-09-062-126-10
; Sequence 10, Application US/09062126
; Patent No. 6500435
; GENERAL INFORMATION:
; APPLICANT: Kamp, Elbarte Margriet
; APPLICANT: Smits, Marinus Adrianus
; TITLE OF INVENTION: Recombinant Vaccine For Prevention
; TITLE OF INVENTION: And/Or Treatment Of Pleuropneumonia Infections
; FILE REFERENCE: 470-980537
; CURRENT APPLICATION NUMBER: US/09/062,126
; CURRENT FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 08/488,706
; PRIOR FILING DATE: 1995-06-09
; PRIOR APPLICATION NUMBER: 08/138,609
; PRIOR FILING DATE: 1993-10-15
; PRIOR APPLICATION NUMBER: 07/722,971
; PRIOR FILING DATE: 1991-06-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; TYPE: PRT
; LENGTH: 1053
; ORGANISM: Actinobacillus pleuropneumonia
; US-09-062-126-10

Query Match 1.5%; Score 14; DB 4; Length 1053;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 KVAAGFELSNOVIG 254
Db 275 KVAAGFELSNOVIG 288

RESULT 6
PCT-US93-10500-2
; Sequence 2, Application PC/TUS9310500
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; TITLE OF INVENTION: Recombinant Vaccine For Procine
; TITLE OF INVENTION: Pleuropneumoniae
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Alan S. Kortman
; STREET: 1600 Empire Tower
; CITY: Buffalo
; STATE: New York
; COUNTRY: U.S.A.
```

```
;
;
; ZIP: 14202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10500
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/972,229
; FILING DATE: 05-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Korman, Alan S.
; REGISTRATION NUMBER: 33,932
; REFERENCE/DOCKET NUMBER: 19603/00001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-853-8104
; TELEFAX: 716-853-8109
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1244 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Actinobacillus pleuropneumoniae
; STRAIN: Serotypes 2, 3, 4, 6 and 8
; INDIVIDUAL ISOLATE: Swine
; CELL TYPE: Gram negative bacterium
; PCT-US93-10500-2

Query Match 1.5%; Score 14; DB 5; Length 1244;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 KVAAGFELSNOVIG 254
Db 448 KVAAGFELSNOVIG 461
|||||

RESULT 7
US-08-772-270A-8
; Sequence 8, Application US/08772270A
; Patent No. 6019984
; GENERAL INFORMATION:
; APPLICANT: MacInnes, Janet
; APPLICANT: Ricciatti, Paul
; APPLICANT: Mallard, Bonnie
; APPLICANT: Rosendal, Soren
; TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
; TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,270A
; FILING DATE: December 23, 1996
```

```
;
;
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 6580-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 956 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Actinobacillus pleuropneumoniae
; US-08-772-270A-8

Query Match 1.2%; Score 11; DB 3; Length 956;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 369 ALLVAGVTGLI 379
Db 392 ALLVAGVTGLI 402
|||||

RESULT 8
US-09-062-126-8
; Sequence 8, Application US/09062126
; Patent No. 6500435
; GENERAL INFORMATION:
; APPLICANT: Kamp, Elbarte Margriet
; APPLICANT: Smits, Marinus Adrianus
; TITLE OF INVENTION: Recombinant Vaccine For Prevention
; TITLE OF INVENTION: And/Or Treatment Of Pleuropneumonia Infections
; FILE REFERENCE: 470-980537
; CURRENT APPLICATION NUMBER: US/09/062,126
; CURRENT FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 08/488,706
; PRIOR FILING DATE: 1995-06-09
; PRIOR APPLICATION NUMBER: 08/138,609
; PRIOR FILING DATE: 1993-10-15
; PRIOR APPLICATION NUMBER: 07/722,971
; PRIOR FILING DATE: 1991-06-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; TYPE: PRT
; LENGTH: 956
; ORGANISM: Actinobacillus pleuropneumonia
; US-09-062-126-8

Query Match 1.2%; Score 11; DB 4; Length 956;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 369 ALLVAGVTGLI 379
Db 392 ALLVAGVTGLI 402
|||||

RESULT 9
US-09-062-126-3
; Sequence 3, Application US/09062126
; Patent No. 6500435
; GENERAL INFORMATION:
; APPLICANT: Kamp, Elbarte Margriet
; APPLICANT: Smits, Marinus Adrianus
; TITLE OF INVENTION: Recombinant Vaccine For Prevention
; TITLE OF INVENTION: And/Or Treatment Of Pleuropneumonia Infections
; FILE REFERENCE: 470-980537
```

; CURRENT APPLICATION NUMBER: US/09/062,126
; CURRENT FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 08/488,706
; PRIOR FILING DATE: 1995-06-09
; PRIOR APPLICATION NUMBER: 08/138,609
; PRIOR FILING DATE: 1993-10-15
; PRIOR APPLICATION NUMBER: 07/722,971
; PRIOR FILING DATE: 1991-06-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Actinobacillus pleuropneumonia
US-09-062-126-3

Query Match 1.2%; Score 11; DB 4; Length 1013;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 266 AQRVAAGLSTT 276
Db 289 AQRVAAGLSTT 299
|||||

RESULT 10
US-08-772-270A-2
; Sequence 2, Application US/08772270A
; Patent No. 6019984
; GENERAL INFORMATION:
; APPLICANT: MacInnes, Janet
; APPLICANT: Ricciatti, Paul
; APPLICANT: Mallard, Bonnie
; APPLICANT: Rosendal, Soren
; TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
; TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,270A
; FILING DATE: December 23, 1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 6580-81
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1022 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Actinobacillus pleuropneumoniae
US-08-772-270A-2

Query Match 1.2%; Score 11; DB 3; Length 1022;
Best Local Similarity 100.0%; Pred. No. 0.18;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 266 AQRVAAGLSTT 276
Db 289 AQRVAAGLSTT 299
|||||

RESULT 11
US-08-387-942C-45
; Sequence 45, Application US/08387942C
; Patent No. 5939289
; GENERAL INFORMATION:
; APPLICANT: ERTESVAG, HELGA
; APPLICANT: VALLA, SVEIN
; APPLICANT: SKJAK-BRAEK, GUDMUND
; APPLICANT: LARSEN, BJORN
; TITLE OF INVENTION: DNA COMPOUNDS COMPRISING SEQUENCES
; TITLE OF INVENTION: ENCODING MANNUROAN C-5-EPIMERASE
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: P.O.BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,942C
; FILING DATE: 09-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1809-106P
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 149 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-387-942C-45

Query Match 1.1%; Score 10; DB 2; Length 149;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 714 GDLLDGGAG 723
Db 31 GDLLDGGAG 40
|||||

RESULT 12
US-09-252-149B-26
; Sequence 26, Application US/09252149B
; Patent No. 6369201
; GENERAL INFORMATION:
; APPLICANT: Barker, Christopher A.
; APPLICANT: Morsey, Mohamad
; TITLE OF INVENTION: IMMUNOLOGICAL METHODS TO MODULAR MYOSTATIN IN
; TITLE OF INVENTION: VERTEBRATE SUBJECTS
; FILE REFERENCE: 9001-0042
; CURRENT APPLICATION NUMBER: US/09/252,149B
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/075,213


```

; PRIOR FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: leukotoxin polypeptide carrier
; OTHER INFORMATION: Figures 15A-15D
US-09-252-149B-26

Query Match          1.1%; Score 10; DB 4; Length 490;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      265 LAQRVAAGLS 274
Db      256 LAQRVAAGLS 265

RESULT 13
US-08-387-156-10
; Sequence 10, Application US/08387156
; Patent No. 5723129
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GARH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,156
; FILING DATE: 10-FEB-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-387-156-10

Query Match          1.1%; Score 10; DB 1; Length 544;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      265 LAQRVAAGLS 274
Db      256 LAQRVAAGLS 265

RESULT 14
US-08-694-865-10
; Sequence 10, Application US/08694865
; Patent No. 5837268
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNIS, JOHN G.
; TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-694-865-10

Query Match          1.1%; Score 10; DB 2; Length 544;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      265 LAQRVAAGLS 274
Db      256 LAQRVAAGLS 265

RESULT 15
US-08-878-748-10
; Sequence 10, Application US/08878748
; Patent No. 5969126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
US-08-878-748-10
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/878,748
FILING DATE: 19-JUN-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-748-10

Query Match 1.1%; Score 10; DB 2; Length 544;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAORVAAGLS 274
Db 256 LAORVAAGLS 265

RESULT 16
US-09-124-491-10
Sequence 10, Application US/09124491
Patent No. 6022960
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANNS, JOHN G.
TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,491
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171

FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-10

Query Match 1.1%; Score 10; DB 3; Length 544;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAORVAAGLS 274
Db 256 LAORVAAGLS 265

RESULT 17
US-09-383-912-10
Sequence 10, Application US/09383912
Patent No. 6521746
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANNS, JOHN G.
TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/383,912
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-383-912-10

Query Match 1.1%; Score 10; DB 4; Length 544;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAORVAAGLS 274

```
Db      256 LAQVAAAGLS 265
|||||
RESULT 18
US-08-694-865-16
; Sequence 16, Application US/08694865
; Patent No. 5837268
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNIS, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-694-865-16

Query Match      1.1%; Score 10; DB 2; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      265 LAQVAAAGLS 274
|||||
Db      359 LAQVAAAGLS 368

RESULT 19
US-09-124-491-16
; Sequence 16, Application US/09124491
; Patent No. 602360
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNIS, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-124-491-16

Query Match      1.1%; Score 10; DB 3; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      265 LAQVAAAGLS 274
|||||
Db      359 LAQVAAAGLS 368

RESULT 20
US-09-383-912-16
; Sequence 16, Application US/09383912
; Patent No. 6521746
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNIS, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 699 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-09-383-912-16

Query Match      1.1%; Score 10; DB 4; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 359 LAQRVAAGLS 368

RESULT 21
US-08-619-812-8
; Sequence 8, Application US/08619812
; Patent No. 6100066
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: THEISEN, MICHAEL
; APPLICANT: HARLAND, RICHARD J.
; APPLICANT: RIOUX, CLEMENT R.
; TITLE OF INVENTION: VACCINES FOR HAEMOPHILUS SOMNUS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,812
; FILING DATE: 15-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/038,719
; FILING DATE: 29-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9000-0019.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 924 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-619-812-8

Query Match      1.1%; Score 10; DB 3; Length 924;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 256 LAQRVAAGLS 265

US-07-908-253-2
; Sequence 2, Application US/07908253
; Patent No. 5534256
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: HARLAND, RICHARD J.
; TITLE OF INVENTION: HAEMOPHILUS SOMNUS OUTER MEMBRANE
; TITLE OF INVENTION: PROTEIN EXTRACT ENRICHED WITH IRON-REGULATED PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBERTA L. ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/908,253
; FILING DATE: 19920702
; CLASSIFICATION: 420
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9000-0026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 926 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-07-908-253-2

Query Match      1.1%; Score 10; DB 1; Length 926;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 256 LAQRVAAGLS 265

RESULT 23
US-08-455-970A-2
; Sequence 2, Application US/08455970A
; Patent No. 5708155
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HOW P.A.
; TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
; TITLE OF INVENTION: CHIMERAS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455.970A
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960.932
; FILING DATE: 14-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-455-970A-2

Query Match 1.1%; Score 10; DB 1; Length 926;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQ RVAAGLS 274
Db 256 LAQ RVAAGLS 265

RESULT 24
US-08-387-156-6
; Sequence 6, Application US/08387156
; Patent No. 5723129
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,156
; FILING DATE: 10-FEB-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 amino acids

;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-387-156-6

Query Match 1.1%; Score 10; DB 1; Length 926;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQ RVAAGLS 274
Db 256 LAQ RVAAGLS 265

RESULT 25
US-08-694-865-6
; Sequence 6, Application US/08694865
; Patent No. 5837268
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNS, JOHN G.
; TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-694-865-6

Query Match 1.1%; Score 10; DB 2; Length 926;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQ RVAAGLS 274
Db 256 LAQ RVAAGLS 265

RESULT 26
US-08-878-748-6
; Sequence 6, Application US/08878748
; Patent No. 5969126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
```

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,748
; FILING DATE: 19-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-878-748-6

Query Match 1.1%; Score 10; DB 2; Length 926;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAORVAAGLS 274
Db 256 LAORVAAGLS 265

RESULT 27
US-08-535-837-2
; Sequence 2, Application US/08535837
; Patent No. 5985289
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: HARLAND, RICHARD J.
; TITLE OF INVENTION: KAEOPHILUS SOMNUS OUTER MEMBRANE
; TITLE OF INVENTION: PROTEIN EXTRACT ENRICHED WITH IRON-REGULATED PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBERTA L. ROBINS
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/535,837
; FILING DATE: 27-SEP-1995
```

```
;
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0026.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-535-837-2

Query Match 1.1%; Score 10; DB 2; Length 926;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAORVAAGLS 274
Db 256 LAORVAAGLS 265

RESULT 28
US-09-124-491-6
; Sequence 6, Application US/09124491
; Patent No. 6022960
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNS, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,491
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE: 09-AUG-1996
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: protein
US-09-124-491-6

Query Match      1.1%; Score 10; DB 3; Length 926;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      265 LAQ RVAAGLS 274
Db      256 LAQ RVAAGLS 265

RESULT 29
US-09-383-912-6
; Sequence 6, Application US/09383912
; Patent No. 6521746
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-383-912-6

Query Match      1.1%; Score 10; DB 4; Length 926;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      265 LAQ RVAAGLS 274
Db      256 LAQ RVAAGLS 265

RESULT 30
5476657-3
; Patent No. 5476657
; APPLICANT: POTTER, ANDREW A.
; TITLE OF INVENTION: PASTEURELLA HAEMOLYTICA LEUKOTOXIN
; COMPOSITIONS AND USES THEREOF
; NUMBER OF SEQUENCES: 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/15,537
```

```
; FILING DATE: 09-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 504,850
; FILING DATE: 05-APR-1990
; APPLICATION NUMBER: 335,018
; FILING DATE: 07-APR-1989
; SEQ ID NO: 3:
; LENGTH: 926
5476657-3

Query Match      1.1%; Score 10; DB 6; Length 926;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      265 LAQ RVAAGLS 274
Db      256 LAQ RVAAGLS 265

RESULT 31
US-08-215-805A-80
; Sequence 80, Application US/08215805A
; Patent No. 5559008
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; TITLE OF INVENTION: LEUKOTOXIN GENE FROM PASTEURELLA
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/215,805A
; FILING DATE: 22-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/61 (D-1329A)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Pasteurella suis
; STRAIN: 5943
; IMMEDIATE SOURCE:
; LIBRARY: P. suis DNA in Bacteriophage lambda-dash
; CLONE: (Lambda)yfc33-37
US-08-215-805A-80

Query Match      1.1%; Score 10; DB 1; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      265 LAQ RVAAGLS 274
Db      279 LAQ RVAAGLS 288
```



```
RESULT 32
US-08-455-970A-12
; Sequence 12, Application US/08455970A
; Patent No. 5708155
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,970A
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 936 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-970A-12

Query Match 1.1%; Score 10; DB 1; Length 936;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 256 LAQRVAAGLS 265

RESULT 33
US-08-455-970A-10
; Sequence 10, Application US/08455970A
; Patent No. 5708155
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,970A
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 936 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-970A-10

Query Match 1.1%; Score 10; DB 1; Length 936;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 256 LAQRVAAGLS 265
```

```
ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,970A
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 943 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-970A-10

Query Match 1.1%; Score 10; DB 1; Length 943;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 256 LAQRVAAGLS 265

RESULT 34
US-08-455-970A-14
; Sequence 14, Application US/08455970A
; Patent No. 5708155
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,970A
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
```

TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 951 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-455-970A-14

Query Match 1.1%; Score 10; DB 1; Length 951;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 256 LAQRVAAGLS 265

RESULT 35
US-08-387-156-8
Sequence 8, Application US/08387156
Patent No. 5723129
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HUW P.A.
TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/387,156
FILING DATE: 10-FEB-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 977 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-387-156-8

Query Match 1.1%; Score 10; DB 1; Length 977;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 256 LAQRVAAGLS 265

RESULT 36
US-08-694-865-8
Sequence 8, Application US/08694865
Patent No. 5837268
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,865
FILING DATE: 09-AUG-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 327-3400
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 977 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-865-8

Query Match 1.1%; Score 10; DB 2; Length 977;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRVAAGLS 274
Db 256 LAQRVAAGLS 265

RESULT 37
US-08-878-748-8
Sequence 8, Application US/08878748
Patent No. 5969126
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HUW P.A.
TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/878,748
; FILING DATE: 19-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 8:
; LENGTH: 977 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-878-748-8

Query Match 1.1%; Score 10; DB 2; Length 977;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQRVAAGLS 274
DB 256 LAQRVAAGLS 265
|||||

RESULT 38
US-09-124-491-8
; Sequence 8, Application US/09124491
; Patent No. 6022960
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,491
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-124-491-8

Query Match 1.1%; Score 10; DB 4; Length 977;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQRVAAGLS 274
DB 256 LAQRVAAGLS 265
|||||

RESULT 39
US-09-383-912-8
; Sequence 8, Application US/09383912
; Patent No. 6521746
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-383-912-8

Query Match 1.1%; Score 10; DB 4; Length 977;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQRVAAGLS 274
DB 256 LAQRVAAGLS 265
|||||

RESULT 38
US-09-124-491-8
; Sequence 8, Application US/09124491
; Patent No. 6022960
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,491
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-124-491-8

Query Match 1.1%; Score 10; DB 2; Length 977;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQRVAAGLS 274
DB 256 LAQRVAAGLS 265
|||||

RESULT 39
US-09-383-912-8
; Sequence 8, Application US/09383912
; Patent No. 6521746
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-383-912-8

Query Match 1.1%; Score 10; DB 4; Length 977;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 LAQRVAAGLS 274
DB 256 LAQRVAAGLS 265
|||||
```

RESULT 40
US-07-777-715-9
; Sequence 9, Application US/07777715
; Patent No. 5273889
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew
; APPLICANT: Campos, Manuel
; APPLICANT: Hughes, Huw P.A.
; TITLE OF INVENTION: CYTOKINE-LEUKOTOXIN GENE FUSIONS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Roerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/777,715
; FILING DATE: 19911016
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Robins, Roberta L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 29310-2001320
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; TELEFAX: 415-327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1069 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-777-715-9

Query Match 1.1%; Score 10; DB 1; Length 1069;
Best Local Similarity 100.0%; Pred.No. 1.6;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 265 LAQRVAAGLS 274
| | | | | | | | | |
Db 256 LAQRVAAGLS 265

Search completed: February 17, 2004, 10:22:10
Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:21:31 ; Search time 39 Seconds
(without alignments)

4976.856 Million cell updates/sec

Title: US-10-069-799-5

Perfect score: 927

Sequence: 1 MSNINIKNIQAGLNSTKS.....SSNALQPIQTPTGILAPSV 927

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 801455 seqs, 209382283 residues

Word size : 0

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 826 | 89.1 | 927 | 11 | US-09-884-696-2 |
| 2 | 14 | 1.5 | 14 | 11 | US-09-884-696-13 |
| 3 | 14 | 1.5 | 1023 | 11 | US-09-884-696-5 |
| 4 | 12 | 1.3 | 12 | 11 | US-09-884-696-6 |
| 5 | 11 | 1.2 | 956 | 11 | US-09-884-696-4 |
| 6 | 10 | 1.1 | 695 | 11 | US-09-305-924-13 |
| 7 | 10 | 1.1 | 953 | 11 | US-09-884-696-3 |
| 8 | 10 | 1.1 | 1017 | 12 | US-10-369-493-20166 |
| 9 | 9 | 1.0 | 259 | 15 | US-10-156-761-12748 |
| 10 | 8 | 0.9 | 307 | 15 | US-10-156-761-12788 |
| 11 | 8 | 0.9 | 408 | 15 | US-10-156-761-14641 |
| 12 | 8 | 0.9 | 424 | 12 | US-10-369-493-17978 |
| 13 | 8 | 0.9 | 477 | 12 | US-10-369-493-18633 |
| 14 | 8 | 0.9 | 489 | 12 | US-10-369-493-16555 |
| 15 | 8 | 0.9 | 505 | 12 | US-10-369-493-17426 |

| | | | | |
|------|-----|----|---------------------|-------------------|
| 527 | 0.9 | 15 | US-10-156-761-11340 | Sequence 11340, A |
| 773 | 0.9 | 12 | US-10-369-493-1471 | Sequence 1471, Ap |
| 4327 | 0.9 | 12 | US-10-369-493-10178 | Sequence 10178, A |
| 20 | 0.8 | 12 | US-10-414-692-80 | Sequence 80, Appl |
| 44 | 0.8 | 8 | US-08-424-550B-519 | Sequence 519, App |
| 49 | 0.8 | 12 | US-10-264-049-3945 | Sequence 3945, Ap |
| 52 | 0.8 | 15 | US-10-185-050-78 | Sequence 78, Appl |
| 63 | 0.8 | 9 | US-09-917-340-9 | Sequence 9, Appli |
| 89 | 0.8 | 9 | US-09-739-907-83 | Sequence 83, Appl |
| 89 | 0.8 | 12 | US-09-938-671-83 | Sequence 83, Appl |
| 90 | 0.8 | 12 | US-09-864-408A-5506 | Sequence 5506, Ap |
| 100 | 0.8 | 12 | US-10-108-260A-3335 | Sequence 3335, Ap |
| 102 | 0.8 | 10 | US-09-924-358-20 | Sequence 20, Appl |
| 102 | 0.8 | 12 | US-10-410-764-20 | Sequence 20, Appl |
| 103 | 0.8 | 12 | US-09-815-242-11108 | Sequence 11108, A |
| 104 | 0.8 | 9 | US-09-741-669-297 | Sequence 297, App |
| 104 | 0.8 | 9 | US-09-912-020-321 | Sequence 321, App |
| 104 | 0.8 | 9 | US-09-815-242-10348 | Sequence 10348, A |
| 104 | 0.8 | 9 | US-09-815-242-14107 | Sequence 14107, A |
| 104 | 0.8 | 12 | US-10-287-274-323 | Sequence 323, App |
| 109 | 0.8 | 10 | US-09-989-919-93 | Sequence 93, Appl |
| 114 | 0.8 | 9 | US-09-864-761-40104 | Sequence 40104, A |
| 115 | 0.8 | 12 | US-10-264-049-3972 | Sequence 3972, Ap |
| 116 | 0.8 | 12 | US-10-094-749-2223 | Sequence 2223, Ap |
| 118 | 0.8 | 9 | US-09-739-907-182 | Sequence 182, App |
| 118 | 0.8 | 12 | US-09-938-671-182 | Sequence 182, App |
| 121 | 0.8 | 9 | US-09-745-763-110 | Sequence 110, App |
| 122 | 0.8 | 12 | US-10-312-187-52 | Sequence 52, Appl |
| 123 | 0.8 | 12 | US-10-386-050A-8 | Sequence 8, Appli |
| 123 | 0.8 | 12 | US-10-386-050A-10 | Sequence 10, Appl |
| 123 | 0.8 | 12 | US-10-386-050A-22 | Sequence 22, Appl |
| 128 | 0.8 | 12 | US-10-291-172-198 | Sequence 198, App |
| 134 | 0.8 | 9 | US-09-925-301-10222 | Sequence 1022, Ap |
| 134 | 0.8 | 10 | US-09-796-692-1206 | Sequence 1206, Ap |
| 134 | 0.8 | 10 | US-09-796-692-1395 | Sequence 1395, Ap |
| 134 | 0.8 | 12 | US-09-796-692-2226 | Sequence 2226, Ap |
| 134 | 0.8 | 12 | US-10-057-475B-1206 | Sequence 1206, Ap |
| 134 | 0.8 | 12 | US-10-057-475B-1985 | Sequence 1985, Ap |
| 134 | 0.8 | 12 | US-10-057-475B-2226 | Sequence 2226, Ap |
| 134 | 0.8 | 12 | US-10-154-884B-1206 | Sequence 1206, Ap |
| 134 | 0.8 | 12 | US-10-154-884B-1985 | Sequence 1985, Ap |
| 134 | 0.8 | 12 | US-10-154-884B-2226 | Sequence 2226, Ap |
| 134 | 0.8 | 15 | US-10-040-862-1206 | Sequence 1206, Ap |
| 134 | 0.8 | 15 | US-10-040-862-1985 | Sequence 1985, Ap |
| 134 | 0.8 | 15 | US-10-040-862-2226 | Sequence 2226, Ap |
| 141 | 0.8 | 14 | US-10-108-605-185 | Sequence 185, App |
| 151 | 0.8 | 12 | US-10-376-564-58 | Sequence 58, Appl |
| 153 | 0.8 | 12 | US-10-376-564-57 | Sequence 57, Appl |
| 155 | 0.8 | 12 | US-10-094-749-2821 | Sequence 2821, Ap |
| 161 | 0.8 | 12 | US-10-369-493-3946 | Sequence 3946, Ap |
| 174 | 0.8 | 12 | US-10-264-237-2789 | Sequence 2789, Ap |
| 187 | 0.8 | 10 | US-09-801-368-136 | Sequence 136, App |
| 187 | 0.8 | 12 | US-10-369-493-22119 | Sequence 22119, A |
| 189 | 0.8 | 9 | US-09-925-301-906 | Sequence 906, App |
| 190 | 0.8 | 10 | US-09-764-868-862 | Sequence 862, App |
| 190 | 0.8 | 12 | US-09-764-875-915 | Sequence 915, App |
| 190 | 0.8 | 12 | US-09-764-875-1191 | Sequence 1191, Ap |
| 203 | 0.8 | 12 | US-10-264-237-1513 | Sequence 1513, Ap |
| 209 | 0.8 | 12 | US-10-427-442-20 | Sequence 20, Appl |
| 219 | 0.8 | 9 | US-09-815-242-11973 | Sequence 11973, A |
| 219 | 0.8 | 15 | US-10-205-823-313 | Sequence 313, App |
| 221 | 0.8 | 12 | US-10-369-493-9552 | Sequence 9552, Ap |
| 235 | 0.8 | 12 | US-10-029-386-32032 | Sequence 32032, A |
| 241 | 0.8 | 9 | US-09-815-242-11653 | Sequence 11653, A |
| 246 | 0.8 | 12 | US-10-369-493-11394 | Sequence 11394, A |
| 250 | 0.8 | 12 | US-10-369-493-14790 | Sequence 14790, A |
| 252 | 0.8 | 12 | US-10-369-493-19889 | Sequence 19889, A |
| 254 | 0.8 | 15 | US-10-156-761-14194 | Sequence 14194, A |
| 256 | 0.8 | 9 | US-09-815-242-10664 | Sequence 10664, A |
| 257 | 0.8 | 11 | US-09-934-900-6 | Sequence 6, Appli |
| 260 | 0.8 | 12 | US-10-156-761-13194 | Sequence 13194, A |
| 269 | 0.8 | 12 | US-10-264-237-2168 | Sequence 2168, Ap |
| 274 | 0.8 | 12 | US-09-953-348-125 | Sequence 125, App |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|----|---------------------|-------------------|-----|---|-----|------|----|---------------------|-------------------|
| 89 | 7 | 0.8 | 274 | 15 | US-10-267-255-125 | Sequence 125, App | 162 | 7 | 0.8 | 457 | 12 | US-10-369-493-9323 | Sequence 9323, Ap |
| 90 | 7 | 0.8 | 275 | 9 | US-09-935-302-737 | Sequence 737, App | 163 | 7 | 0.8 | 458 | 9 | US-09-815-242-11321 | Sequence 11321, A |
| 91 | 7 | 0.8 | 283 | 12 | US-09-930-512-14 | Sequence 14, Appl | 164 | 7 | 0.8 | 460 | 11 | US-09-759-595-16 | Sequence 11, Appl |
| 92 | 7 | 0.8 | 284 | 12 | US-10-114-828-18 | Sequence 18, Appl | 165 | 7 | 0.8 | 460 | 15 | US-10-001-254-16 | Sequence 16, Appl |
| 93 | 7 | 0.8 | 296 | 12 | US-10-096-494-1 | Sequence 1, Appl | 166 | 7 | 0.8 | 460 | 15 | US-10-001-254-28 | Sequence 28, Appl |
| 94 | 7 | 0.8 | 296 | 12 | US-10-369-493-8679 | Sequence 8679, Ap | 167 | 7 | 0.8 | 463 | 15 | US-10-156-761-14014 | Sequence 14014, A |
| 95 | 7 | 0.8 | 303 | 15 | US-10-156-761-8055 | Sequence 8055, Ap | 168 | 7 | 0.8 | 468 | 11 | US-09-746-660A-66 | Sequence 66, Appl |
| 96 | 7 | 0.8 | 308 | 10 | US-09-738-626-6317 | Sequence 6317, Ap | 169 | 7 | 0.8 | 468 | 12 | US-10-354-774-46 | Sequence 46, Appl |
| 97 | 7 | 0.8 | 313 | 12 | US-10-306-762-59 | Sequence 59, Appl | 170 | 7 | 0.8 | 472 | 12 | US-10-271-012-46 | Sequence 46, Appl |
| 98 | 7 | 0.8 | 320 | 9 | US-09-854-122-20 | Sequence 20, Appl | 171 | 7 | 0.8 | 472 | 12 | US-10-369-493-17633 | Sequence 17633, A |
| 99 | 7 | 0.8 | 322 | 10 | US-09-905-176-24 | Sequence 24, Appl | 172 | 7 | 0.8 | 478 | 15 | US-10-156-761-11009 | Sequence 11009, A |
| 100 | 7 | 0.8 | 324 | 9 | US-09-815-242-11879 | Sequence 11879, A | 173 | 7 | 0.8 | 481 | 12 | US-10-369-493-10573 | Sequence 10573, A |
| 101 | 7 | 0.8 | 329 | 11 | US-09-950-041-16 | Sequence 16, Appl | 174 | 7 | 0.8 | 493 | 12 | US-10-369-493-23665 | Sequence 23665, A |
| 102 | 7 | 0.8 | 333 | 10 | US-09-976-059-2 | Sequence 2, Appl | 175 | 7 | 0.8 | 494 | 12 | US-10-369-493-12924 | Sequence 12924, A |
| 103 | 7 | 0.8 | 333 | 12 | US-10-438-537-2 | Sequence 2, Appl | 176 | 7 | 0.8 | 494 | 12 | US-10-369-493-20175 | Sequence 20175, A |
| 104 | 7 | 0.8 | 338 | 12 | US-10-369-493-16948 | Sequence 16948, A | 177 | 7 | 0.8 | 516 | 12 | US-10-369-493-16571 | Sequence 16571, A |
| 105 | 7 | 0.8 | 339 | 9 | US-09-838-955-8 | Sequence 8, Appl | 178 | 7 | 0.8 | 519 | 15 | US-10-205-823-299 | Sequence 299, App |
| 106 | 7 | 0.8 | 339 | 15 | US-10-156-761-9210 | Sequence 9210, Ap | 179 | 7 | 0.8 | 521 | 10 | US-09-884-566-2 | Sequence 2, Appl |
| 107 | 7 | 0.8 | 339 | 15 | US-10-156-761-12600 | Sequence 12600, A | 180 | 7 | 0.8 | 521 | 15 | US-10-156-761-12853 | Sequence 12853, A |
| 108 | 7 | 0.8 | 341 | 12 | US-10-369-493-18432 | Sequence 18432, A | 181 | 7 | 0.8 | 523 | 15 | US-10-219-220-183 | Sequence 183, App |
| 109 | 7 | 0.8 | 342 | 16 | US-10-233-926-18 | Sequence 18, Appl | 182 | 7 | 0.8 | 551 | 10 | US-09-738-626-431 | Sequence 431, Ap |
| 110 | 7 | 0.8 | 343 | 15 | US-10-156-761-9520 | Sequence 9520, Ap | 183 | 7 | 0.8 | 555 | 12 | US-10-369-493-17959 | Sequence 17959, A |
| 111 | 7 | 0.8 | 349 | 16 | US-10-233-926-2 | Sequence 2, Appl | 184 | 7 | 0.8 | 585 | 9 | US-09-815-242-11132 | Sequence 11132, A |
| 112 | 7 | 0.8 | 349 | 16 | US-10-233-926-12 | Sequence 12, Appl | 185 | 7 | 0.8 | 587 | 10 | US-09-815-923-4 | Sequence 4, Appl |
| 113 | 7 | 0.8 | 350 | 12 | US-10-369-493-20851 | Sequence 20851, A | 186 | 7 | 0.8 | 587 | 12 | US-10-238-075-1121 | Sequence 1121, Ap |
| 114 | 7 | 0.8 | 353 | 15 | US-10-106-698-6180 | Sequence 6180, Ap | 187 | 7 | 0.8 | 587 | 10 | US-10-320-800-54 | Sequence 54, Appl |
| 115 | 7 | 0.8 | 358 | 16 | US-10-080-170-440 | Sequence 440, App | 188 | 7 | 0.8 | 595 | 12 | US-10-369-493-5268 | Sequence 5268, Ap |
| 116 | 7 | 0.8 | 359 | 12 | US-10-369-493-20613 | Sequence 20613, A | 189 | 7 | 0.8 | 600 | 12 | US-10-369-493-18505 | Sequence 18505, A |
| 117 | 7 | 0.8 | 361 | 12 | US-10-369-493-9259 | Sequence 9259, Ap | 190 | 7 | 0.8 | 605 | 12 | US-10-369-493-19505 | Sequence 19505, A |
| 118 | 7 | 0.8 | 361 | 12 | US-10-369-493-9446 | Sequence 9446, Ap | 191 | 7 | 0.8 | 605 | 12 | US-10-104-047-3739 | Sequence 3739, Ap |
| 119 | 7 | 0.8 | 373 | 12 | US-10-264-049-3049 | Sequence 3049, Ap | 192 | 7 | 0.8 | 617 | 12 | US-10-354-358-118 | Sequence 118, App |
| 120 | 7 | 0.8 | 374 | 12 | US-10-369-493-17585 | Sequence 17585, A | 193 | 7 | 0.8 | 622 | 9 | US-09-843-598-11 | Sequence 11, Appl |
| 121 | 7 | 0.8 | 379 | 12 | US-10-158-034-65 | Sequence 65, Appl | 194 | 7 | 0.8 | 622 | 12 | US-10-369-493-21226 | Sequence 21226, A |
| 122 | 7 | 0.8 | 379 | 12 | US-10-158-034-90 | Sequence 90, Appl | 195 | 7 | 0.8 | 630 | 12 | US-10-369-493-421 | Sequence 421, App |
| 123 | 7 | 0.8 | 380 | 11 | US-09-934-900-12 | Sequence 12, Appl | 196 | 7 | 0.8 | 630 | 12 | US-10-369-493-14775 | Sequence 14775, A |
| 124 | 7 | 0.8 | 380 | 12 | US-10-411-976-9 | Sequence 9, Appl | 197 | 7 | 0.8 | 634 | 12 | US-10-369-493-11753 | Sequence 11753, A |
| 125 | 7 | 0.8 | 382 | 10 | US-09-881-752A-224 | Sequence 224, App | 198 | 7 | 0.8 | 635 | 12 | US-10-369-493-14237 | Sequence 14237, A |
| 126 | 7 | 0.8 | 382 | 14 | US-10-078-929-34 | Sequence 34, Appl | 199 | 7 | 0.8 | 636 | 12 | US-10-369-493-14928 | Sequence 14928, A |
| 127 | 7 | 0.8 | 385 | 15 | US-10-156-761-12542 | Sequence 12542, A | 200 | 7 | 0.8 | 637 | 12 | US-10-108-260A-4790 | Sequence 4790, Ap |
| 128 | 7 | 0.8 | 386 | 10 | US-09-881-752A-74 | Sequence 74, Appl | 201 | 7 | 0.8 | 640 | 12 | US-10-369-493-12290 | Sequence 12290, A |
| 129 | 7 | 0.8 | 386 | 11 | US-09-988-0678-32 | Sequence 32, Appl | 202 | 7 | 0.8 | 641 | 12 | US-10-369-493-20270 | Sequence 20270, A |
| 130 | 7 | 0.8 | 386 | 12 | US-10-369-493-12670 | Sequence 12670, A | 203 | 7 | 0.8 | 655 | 12 | US-10-262-666-8 | Sequence 8, Appl |
| 131 | 7 | 0.8 | 386 | 15 | US-10-204-887-142 | Sequence 142, App | 204 | 7 | 0.8 | 660 | 9 | US-09-841-132-446 | Sequence 446, App |
| 132 | 7 | 0.8 | 388 | 12 | US-10-369-493-19947 | Sequence 19947, A | 205 | 7 | 0.8 | 662 | 12 | US-10-369-493-5336 | Sequence 5336, Ap |
| 133 | 7 | 0.8 | 389 | 12 | US-10-369-493-19972 | Sequence 19972, A | 206 | 7 | 0.8 | 674 | 12 | US-10-291-265-471 | Sequence 471, App |
| 134 | 7 | 0.8 | 389 | 14 | US-10-036-507-13 | Sequence 13, Appl | 207 | 7 | 0.8 | 688 | 12 | US-10-369-493-20655 | Sequence 20655, A |
| 135 | 7 | 0.8 | 390 | 12 | US-10-369-493-18060 | Sequence 18060, A | 208 | 7 | 0.8 | 708 | 12 | US-10-369-493-15900 | Sequence 15900, A |
| 136 | 7 | 0.8 | 390 | 12 | US-10-369-493-19694 | Sequence 19694, A | 209 | 7 | 0.8 | 708 | 12 | US-10-369-493-16271 | Sequence 16271, A |
| 137 | 7 | 0.8 | 393 | 12 | US-10-369-493-20721 | Sequence 20721, A | 210 | 7 | 0.8 | 710 | 9 | US-09-815-242-5367 | Sequence 5367, Ap |
| 138 | 7 | 0.8 | 393 | 12 | US-10-374-780A-1937 | Sequence 1937, Ap | 211 | 7 | 0.8 | 711 | 9 | US-09-815-242-12190 | Sequence 12190, A |
| 139 | 7 | 0.8 | 396 | 12 | US-10-084-846A-76 | Sequence 76, Appl | 212 | 7 | 0.8 | 714 | 15 | US-10-156-761-11928 | Sequence 11928, A |
| 140 | 7 | 0.8 | 400 | 10 | US-09-373-967-2 | Sequence 2, Appl | 213 | 7 | 0.8 | 732 | 11 | US-09-977-418-20 | Sequence 20, Appl |
| 141 | 7 | 0.8 | 400 | 12 | US-10-369-493-17487 | Sequence 17487, A | 214 | 7 | 0.8 | 732 | 11 | US-09-977-033A-20 | Sequence 20, Appl |
| 142 | 7 | 0.8 | 407 | 12 | US-10-166-225A-84 | Sequence 84, Appl | 215 | 7 | 0.8 | 732 | 12 | US-09-977-751C-20 | Sequence 20, Appl |
| 143 | 7 | 0.8 | 424 | 10 | US-09-843-905A-13 | Sequence 13, Appl | 216 | 7 | 0.8 | 732 | 12 | US-09-977-639A-20 | Sequence 20, Appl |
| 144 | 7 | 0.8 | 424 | 12 | US-10-317-250-13 | Sequence 13, Appl | 217 | 7 | 0.8 | 732 | 12 | US-09-977-819B-20 | Sequence 20, Appl |
| 145 | 7 | 0.8 | 435 | 12 | US-09-813-432-48 | Sequence 48, Appl | 218 | 7 | 0.8 | 735 | 12 | US-10-262-666-66 | Sequence 66, Appl |
| 146 | 7 | 0.8 | 435 | 12 | US-10-174-364-48 | Sequence 48, Appl | 219 | 7 | 0.8 | 744 | 12 | US-10-262-666-64 | Sequence 64, Appl |
| 147 | 7 | 0.8 | 436 | 12 | US-10-369-493-21209 | Sequence 21209, A | 220 | 7 | 0.8 | 756 | 12 | US-10-369-493-16891 | Sequence 16891, A |
| 148 | 7 | 0.8 | 440 | 11 | US-09-910-186A-8 | Sequence 8, Appl | 221 | 7 | 0.8 | 781 | 12 | US-10-187-253A-33 | Sequence 33, Appl |
| 149 | 7 | 0.8 | 441 | 12 | US-10-452-024-188 | Sequence 188, App | 222 | 7 | 0.8 | 829 | 12 | US-10-161-927-86 | Sequence 86, Appl |
| 150 | 7 | 0.8 | 442 | 12 | US-10-369-493-20368 | Sequence 20368, A | 223 | 7 | 0.8 | 808 | 15 | US-10-156-761-10993 | Sequence 10993, A |
| 151 | 7 | 0.8 | 442 | 12 | US-10-369-493-2938 | Sequence 2938, Ap | 224 | 7 | 0.8 | 810 | 15 | US-10-156-761-9208 | Sequence 9208, Ap |
| 152 | 7 | 0.8 | 449 | 12 | US-10-369-493-21458 | Sequence 21458, A | 225 | 7 | 0.8 | 812 | 15 | US-10-210-296-6 | Sequence 6, Appl |
| 153 | 7 | 0.8 | 450 | 9 | US-09-795-693-34 | Sequence 34, Appl | 226 | 7 | 0.8 | 827 | 12 | US-10-369-493-21772 | Sequence 21772, A |
| 154 | 7 | 0.8 | 450 | 12 | US-10-302-267-96 | Sequence 96, Appl | 227 | 7 | 0.8 | 828 | 12 | US-10-425-913-3 | Sequence 3, Appl |
| 155 | 7 | 0.8 | 450 | 12 | US-10-374-780A-2128 | Sequence 2128, Ap | 228 | 7 | 0.8 | 847 | 12 | US-10-161-927-86 | Sequence 86, Appl |
| 156 | 7 | 0.8 | 450 | 15 | US-10-156-239-34 | Sequence 34, Appl | 229 | 7 | 0.8 | 847 | 12 | US-10-369-493-6478 | Sequence 6478, Ap |
| 157 | 7 | 0.8 | 450 | 15 | US-10-139-485-34 | Sequence 34, Appl | 230 | 7 | 0.8 | 864 | 12 | US-10-369-493-16631 | Sequence 16631, A |
| 158 | 7 | 0.8 | 454 | 12 | US-10-238-075-1220 | Sequence 1220, Ap | 231 | 7 | 0.8 | 893 | 9 | US-09-916-790-5 | Sequence 5, Appl |
| 159 | 7 | 0.8 | 454 | 12 | US-10-369-493-1302 | Sequence 1302, Ap | 232 | 7 | 0.8 | 907 | 12 | US-10-369-493-15529 | Sequence 15529, A |
| 160 | 7 | 0.8 | 454 | 12 | US-10-369-493-21598 | Sequence 21598, A | 233 | 7 | 0.8 | 1000 | 11 | US-10-823-187-6 | Sequence 6, Appl |
| 161 | 7 | 0.8 | 455 | 9 | US-09-815-242-11558 | Sequence 11558, A | 234 | 7 | 0.8 | 1032 | 11 | US-09-950-041-37 | Sequence 37, Appl |

| | | | | | | | | | | | | | |
|-----|---|-----|------|----|--------------------|-------------------|-----|---|-----|------|----|--------------------|-------------------|
| 527 | 7 | 0.8 | 1036 | 15 | US-10-123-261-142 | Sequence 142, App | 600 | 7 | 0.8 | 1036 | 15 | US-10-175-735-142 | Sequence 142, App |
| 528 | 7 | 0.8 | 1036 | 15 | US-10-140-921-142 | Sequence 142, App | 601 | 7 | 0.8 | 1036 | 15 | US-10-121-040-142 | Sequence 142, App |
| 529 | 7 | 0.8 | 1036 | 15 | US-10-140-928-142 | Sequence 142, App | 602 | 7 | 0.8 | 1036 | 15 | US-10-121-056-142 | Sequence 142, App |
| 530 | 7 | 0.8 | 1036 | 15 | US-10-121-045-142 | Sequence 142, App | 603 | 7 | 0.8 | 1036 | 15 | US-10-121-061-142 | Sequence 142, App |
| 531 | 7 | 0.8 | 1036 | 15 | US-10-123-292-142 | Sequence 142, App | 604 | 7 | 0.8 | 1036 | 15 | US-10-123-235-142 | Sequence 142, App |
| 532 | 7 | 0.8 | 1036 | 15 | US-10-123-903-142 | Sequence 142, App | 605 | 7 | 0.8 | 1036 | 15 | US-10-124-818-142 | Sequence 142, App |
| 533 | 7 | 0.8 | 1036 | 15 | US-10-124-819-142 | Sequence 142, App | 606 | 7 | 0.8 | 1036 | 15 | US-10-137-868-142 | Sequence 142, App |
| 534 | 7 | 0.8 | 1036 | 15 | US-10-124-822-142 | Sequence 142, App | 607 | 7 | 0.8 | 1036 | 15 | US-10-147-492-142 | Sequence 142, App |
| 535 | 7 | 0.8 | 1036 | 15 | US-10-140-925-142 | Sequence 142, App | 608 | 7 | 0.8 | 1036 | 15 | US-10-158-782-142 | Sequence 142, App |
| 536 | 7 | 0.8 | 1036 | 15 | US-10-160-498-142 | Sequence 142, App | 609 | 7 | 0.8 | 1036 | 15 | US-10-123-905-142 | Sequence 142, App |
| 537 | 7 | 0.8 | 1036 | 15 | US-10-124-824-142 | Sequence 142, App | 610 | 7 | 0.8 | 1036 | 15 | US-10-123-907-142 | Sequence 142, App |
| 538 | 7 | 0.8 | 1036 | 15 | US-10-127-825A-142 | Sequence 142, App | 611 | 7 | 0.8 | 1036 | 15 | US-10-124-815-142 | Sequence 142, App |
| 539 | 7 | 0.8 | 1036 | 15 | US-10-127-829A-142 | Sequence 142, App | 612 | 7 | 0.8 | 1036 | 15 | US-10-125-921A-142 | Sequence 142, App |
| 540 | 7 | 0.8 | 1036 | 15 | US-10-127-835A-142 | Sequence 142, App | 613 | 7 | 0.8 | 1036 | 15 | US-10-125-928A-142 | Sequence 142, App |
| 541 | 7 | 0.8 | 1036 | 15 | US-10-127-839A-142 | Sequence 142, App | 614 | 7 | 0.8 | 1036 | 15 | US-10-127-821A-142 | Sequence 142, App |
| 542 | 7 | 0.8 | 1036 | 15 | US-10-127-901A-142 | Sequence 142, App | 615 | 7 | 0.8 | 1036 | 15 | US-10-127-822A-142 | Sequence 142, App |
| 543 | 7 | 0.8 | 1036 | 15 | US-10-128-693A-142 | Sequence 142, App | 616 | 7 | 0.8 | 1036 | 15 | US-10-127-824A-142 | Sequence 142, App |
| 544 | 7 | 0.8 | 1036 | 15 | US-10-131-813A-142 | Sequence 142, App | 617 | 7 | 0.8 | 1036 | 15 | US-10-127-826A-142 | Sequence 142, App |
| 545 | 7 | 0.8 | 1036 | 15 | US-10-131-818A-142 | Sequence 142, App | 618 | 7 | 0.8 | 1036 | 15 | US-10-127-827A-142 | Sequence 142, App |
| 546 | 7 | 0.8 | 1036 | 15 | US-10-131-823A-142 | Sequence 142, App | 619 | 7 | 0.8 | 1036 | 15 | US-10-127-828A-142 | Sequence 142, App |
| 547 | 7 | 0.8 | 1036 | 15 | US-10-131-824A-142 | Sequence 142, App | 620 | 7 | 0.8 | 1036 | 15 | US-10-127-830A-142 | Sequence 142, App |
| 548 | 7 | 0.8 | 1036 | 15 | US-10-131-830A-142 | Sequence 142, App | 621 | 7 | 0.8 | 1036 | 15 | US-10-127-832A-142 | Sequence 142, App |
| 549 | 7 | 0.8 | 1036 | 15 | US-10-131-837A-142 | Sequence 142, App | 622 | 7 | 0.8 | 1036 | 15 | US-10-127-833A-142 | Sequence 142, App |
| 550 | 7 | 0.8 | 1036 | 15 | US-10-137-872A-142 | Sequence 142, App | 623 | 7 | 0.8 | 1036 | 15 | US-10-127-834A-142 | Sequence 142, App |
| 551 | 7 | 0.8 | 1036 | 15 | US-10-147-500-142 | Sequence 142, App | 624 | 7 | 0.8 | 1036 | 15 | US-10-127-836A-142 | Sequence 142, App |
| 552 | 7 | 0.8 | 1036 | 15 | US-10-147-502-142 | Sequence 142, App | 625 | 7 | 0.8 | 1036 | 15 | US-10-127-841A-142 | Sequence 142, App |
| 553 | 7 | 0.8 | 1036 | 15 | US-10-147-515-142 | Sequence 142, App | 626 | 7 | 0.8 | 1036 | 15 | US-10-127-844A-142 | Sequence 142, App |
| 554 | 7 | 0.8 | 1036 | 15 | US-10-147-517-142 | Sequence 142, App | 627 | 7 | 0.8 | 1036 | 15 | US-10-128-687A-142 | Sequence 142, App |
| 555 | 7 | 0.8 | 1036 | 15 | US-10-147-526-142 | Sequence 142, App | 628 | 7 | 0.8 | 1036 | 15 | US-10-128-688A-142 | Sequence 142, App |
| 556 | 7 | 0.8 | 1036 | 15 | US-10-147-527-142 | Sequence 142, App | 629 | 7 | 0.8 | 1036 | 15 | US-10-128-689A-142 | Sequence 142, App |
| 557 | 7 | 0.8 | 1036 | 15 | US-10-121-041-142 | Sequence 142, App | 630 | 7 | 0.8 | 1036 | 15 | US-10-128-694A-142 | Sequence 142, App |
| 558 | 7 | 0.8 | 1036 | 15 | US-10-121-043-142 | Sequence 142, App | 631 | 7 | 0.8 | 1036 | 15 | US-10-131-825A-142 | Sequence 142, App |
| 559 | 7 | 0.8 | 1036 | 15 | US-10-121-047-142 | Sequence 142, App | 632 | 7 | 0.8 | 1036 | 15 | US-10-230-417-142 | Sequence 142, App |
| 560 | 7 | 0.8 | 1036 | 15 | US-10-123-215-142 | Sequence 142, App | 633 | 7 | 0.8 | 1036 | 15 | US-10-131-815A-142 | Sequence 142, App |
| 561 | 7 | 0.8 | 1036 | 15 | US-10-123-902-142 | Sequence 142, App | 634 | 7 | 0.8 | 1036 | 15 | US-10-131-817A-142 | Sequence 142, App |
| 562 | 7 | 0.8 | 1036 | 15 | US-10-123-908-142 | Sequence 142, App | 635 | 7 | 0.8 | 1036 | 15 | US-10-131-821A-142 | Sequence 142, App |
| 563 | 7 | 0.8 | 1036 | 15 | US-10-123-909-142 | Sequence 142, App | 636 | 7 | 0.8 | 1036 | 15 | US-10-131-822A-142 | Sequence 142, App |
| 564 | 7 | 0.8 | 1036 | 15 | US-10-123-910-142 | Sequence 142, App | 637 | 7 | 0.8 | 1036 | 15 | US-10-131-828A-142 | Sequence 142, App |
| 565 | 7 | 0.8 | 1036 | 15 | US-10-124-817-142 | Sequence 142, App | 638 | 7 | 0.8 | 1036 | 15 | US-10-131-835A-142 | Sequence 142, App |
| 566 | 7 | 0.8 | 1036 | 15 | US-10-124-818-142 | Sequence 142, App | 639 | 7 | 0.8 | 1036 | 15 | US-10-137-864A-142 | Sequence 142, App |
| 567 | 7 | 0.8 | 1036 | 15 | US-10-125-922-142 | Sequence 142, App | 640 | 7 | 0.8 | 1036 | 15 | US-10-137-869A-142 | Sequence 142, App |
| 568 | 7 | 0.8 | 1036 | 15 | US-10-125-924-142 | Sequence 142, App | 641 | 7 | 0.8 | 1036 | 15 | US-10-147-523-142 | Sequence 142, App |
| 569 | 7 | 0.8 | 1036 | 15 | US-10-140-860-142 | Sequence 142, App | 642 | 7 | 0.8 | 1036 | 15 | US-10-158-785-142 | Sequence 142, App |
| 570 | 7 | 0.8 | 1036 | 15 | US-10-142-417-142 | Sequence 142, App | 643 | 7 | 0.8 | 1036 | 15 | US-10-121-051-142 | Sequence 142, App |
| 571 | 7 | 0.8 | 1036 | 15 | US-10-147-519-142 | Sequence 142, App | 644 | 7 | 0.8 | 1036 | 15 | US-10-121-042-142 | Sequence 142, App |
| 572 | 7 | 0.8 | 1036 | 15 | US-10-157-782-142 | Sequence 142, App | 645 | 7 | 0.8 | 1036 | 15 | US-10-123-912-142 | Sequence 142, App |
| 573 | 7 | 0.8 | 1036 | 15 | US-10-152-395-142 | Sequence 142, App | 646 | 7 | 0.8 | 1036 | 15 | US-10-192-007-142 | Sequence 142, App |
| 574 | 7 | 0.8 | 1036 | 15 | US-10-125-926A-142 | Sequence 142, App | 647 | 7 | 0.8 | 1036 | 15 | US-10-194-359-142 | Sequence 142, App |
| 575 | 7 | 0.8 | 1036 | 15 | US-10-125-930A-142 | Sequence 142, App | 648 | 7 | 0.8 | 1036 | 15 | US-10-127-847A-142 | Sequence 142, App |
| 576 | 7 | 0.8 | 1036 | 15 | US-10-127-831A-142 | Sequence 142, App | 649 | 7 | 0.8 | 1036 | 15 | US-10-175-590-142 | Sequence 142, App |
| 577 | 7 | 0.8 | 1036 | 15 | US-10-127-837A-142 | Sequence 142, App | 650 | 7 | 0.8 | 1036 | 16 | US-10-137-866-142 | Sequence 142, App |
| 578 | 7 | 0.8 | 1036 | 15 | US-10-127-838B-142 | Sequence 142, App | 651 | 7 | 0.8 | 1036 | 16 | US-10-146-726-142 | Sequence 142, App |
| 579 | 7 | 0.8 | 1036 | 15 | US-10-127-842A-142 | Sequence 142, App | 652 | 7 | 0.8 | 1036 | 16 | US-10-146-727-142 | Sequence 142, App |
| 580 | 7 | 0.8 | 1036 | 15 | US-10-127-843A-142 | Sequence 142, App | 653 | 7 | 0.8 | 1036 | 16 | US-10-146-788-142 | Sequence 142, App |
| 581 | 7 | 0.8 | 1036 | 15 | US-10-127-845A-142 | Sequence 142, App | 654 | 7 | 0.8 | 1036 | 16 | US-10-152-380-142 | Sequence 142, App |
| 582 | 7 | 0.8 | 1036 | 15 | US-10-127-846A-142 | Sequence 142, App | 655 | 7 | 0.8 | 1036 | 16 | US-10-153-934-142 | Sequence 142, App |
| 583 | 7 | 0.8 | 1036 | 15 | US-10-127-848A-142 | Sequence 142, App | 656 | 7 | 0.8 | 1040 | 9 | US-09-864-761-3825 | Sequence 3825, A |
| 584 | 7 | 0.8 | 1036 | 15 | US-10-127-849A-142 | Sequence 142, App | 657 | 7 | 0.8 | 1041 | 10 | US-09-978-295A-498 | Sequence 498, App |
| 585 | 7 | 0.8 | 1036 | 15 | US-10-127-850A-142 | Sequence 142, App | 658 | 7 | 0.8 | 1041 | 10 | US-09-978-697-498 | Sequence 498, App |
| 586 | 7 | 0.8 | 1036 | 15 | US-10-127-851A-142 | Sequence 142, App | 659 | 7 | 0.8 | 1041 | 10 | US-09-978-192A-498 | Sequence 498, App |
| 587 | 7 | 0.8 | 1036 | 15 | US-10-128-684A-142 | Sequence 142, App | 660 | 7 | 0.8 | 1041 | 11 | US-09-999-832A-498 | Sequence 498, App |
| 588 | 7 | 0.8 | 1036 | 15 | US-10-128-686A-142 | Sequence 142, App | 661 | 7 | 0.8 | 1041 | 11 | US-09-978-189-498 | Sequence 498, App |
| 589 | 7 | 0.8 | 1036 | 15 | US-10-128-690A-142 | Sequence 142, App | 662 | 7 | 0.8 | 1041 | 11 | US-09-978-608A-498 | Sequence 498, App |
| 590 | 7 | 0.8 | 1036 | 15 | US-10-128-691A-142 | Sequence 142, App | 663 | 7 | 0.8 | 1041 | 11 | US-09-978-585A-498 | Sequence 498, App |
| 591 | 7 | 0.8 | 1036 | 15 | US-10-131-819A-142 | Sequence 142, App | 664 | 7 | 0.8 | 1041 | 11 | US-09-978-191A-498 | Sequence 498, App |
| 592 | 7 | 0.8 | 1036 | 15 | US-10-131-829A-142 | Sequence 142, App | 665 | 7 | 0.8 | 1041 | 11 | US-09-978-403A-498 | Sequence 498, App |
| 593 | 7 | 0.8 | 1036 | 15 | US-10-131-836A-142 | Sequence 142, App | 666 | 7 | 0.8 | 1041 | 11 | US-09-978-564A-498 | Sequence 498, App |
| 594 | 7 | 0.8 | 1036 | 15 | US-10-146-729-142 | Sequence 142, App | 667 | 7 | 0.8 | 1041 | 11 | US-09-999-833A-498 | Sequence 498, App |
| 595 | 7 | 0.8 | 1036 | 15 | US-10-146-791-142 | Sequence 142, App | 668 | 7 | 0.8 | 1041 | 11 | US-09-981-915A-498 | Sequence 498, App |
| 596 | 7 | 0.8 | 1036 | 15 | US-10-147-408-142 | Sequence 142, App | 669 | 7 | 0.8 | 1041 | 11 | US-09-978-824-498 | Sequence 498, App |
| 597 | 7 | 0.8 | 1036 | 15 | US-10-147-508-142 | Sequence 142, App | 670 | 7 | 0.8 | 1041 | 11 | US-09-918-585A-498 | Sequence 498, App |
| 598 | 7 | 0.8 | 1036 | 15 | US-10-147-512-142 | Sequence 142, App | 671 | 7 | 0.8 | 1041 | 11 | US-09-978-423A-498 | Sequence 498, App |
| 599 | 7 | 0.8 | 1036 | 15 | US-10-152-724A-2 | Sequence 2, App11 | 672 | 7 | 0.8 | 1041 | 11 | US-09-978-193A-498 | Sequence 498, App |

| | | | | | | | | | | | | | |
|-----|---|-----|------|----|---------------------|--------------------|-----|---|-----|-------|----|---------------------|--------------------|
| 673 | 7 | 0.8 | 1041 | 11 | US-09-999-830A-498 | Sequence 498, App | 746 | 7 | 0.8 | 1290 | 10 | US-09-881-752A-138 | Sequence 138, App |
| 674 | 7 | 0.8 | 1041 | 11 | US-09-978-757A-498 | Sequence 498, App | 747 | 7 | 0.8 | 1290 | 12 | US-10-360-101-220 | Sequence 220, App |
| 675 | 7 | 0.8 | 1041 | 11 | US-09-978-187B-498 | Sequence 498, App | 748 | 7 | 0.8 | 1290 | 12 | US-10-452-024-119 | Sequence 119, App |
| 676 | 7 | 0.8 | 1041 | 11 | US-09-954-987B-184 | Sequence 184, App | 749 | 7 | 0.8 | 1291 | 12 | US-10-354-774-42 | Sequence 42, Appl |
| 677 | 7 | 0.8 | 1041 | 11 | US-09-954-987B-186 | Sequence 186, App | 750 | 7 | 0.8 | 1291 | 12 | US-10-271-012-42 | Sequence 42, Appl |
| 678 | 7 | 0.8 | 1041 | 11 | US-09-978-643A-498 | Sequence 498, App | 751 | 7 | 0.8 | 1291 | 12 | US-10-452-024-2 | Sequence 2, Appl |
| 679 | 7 | 0.8 | 1041 | 11 | US-09-978-375A-498 | Sequence 498, App | 752 | 7 | 0.8 | 1291 | 12 | US-10-452-024-118 | Sequence 118, App |
| 680 | 7 | 0.8 | 1041 | 12 | US-09-978-188A-498 | Sequence 498, App | 753 | 7 | 0.8 | 1291 | 12 | US-10-452-024-121 | Sequence 121, App |
| 681 | 7 | 0.8 | 1041 | 12 | US-09-978-188A-498 | Sequence 498, App | 754 | 7 | 0.8 | 1291 | 12 | US-10-452-024-122 | Sequence 122, App |
| 682 | 7 | 0.8 | 1041 | 12 | US-09-978-298A-498 | Sequence 498, App | 755 | 7 | 0.8 | 1291 | 12 | US-10-452-024-123 | Sequence 123, App |
| 683 | 7 | 0.8 | 1041 | 12 | US-10-143-031A-498 | Sequence 26, Appl | 756 | 7 | 0.8 | 1291 | 12 | US-10-452-024-142 | Sequence 142, App |
| 684 | 7 | 0.8 | 1041 | 12 | US-10-272-502A-26 | Sequence 498, App | 757 | 7 | 0.8 | 1291 | 15 | US-10-156-761-14161 | Sequence 14161, A |
| 685 | 7 | 0.8 | 1041 | 12 | US-10-002-967A-498 | Sequence 498, App | 758 | 7 | 0.8 | 1435 | 15 | US-10-128-714-8125 | Sequence 8125, Ap |
| 686 | 7 | 0.8 | 1041 | 12 | US-10-017-083A-498 | Sequence 498, App | 759 | 7 | 0.8 | 1908 | 15 | US-10-128-714-3475 | Sequence 3475, Ap |
| 687 | 7 | 0.8 | 1041 | 12 | US-10-143-030A-498 | Sequence 498, App | 760 | 7 | 0.8 | 2022 | 15 | US-10-128-714-8475 | Sequence 8475, Ap |
| 688 | 7 | 0.8 | 1041 | 12 | US-10-145-128A-498 | Sequence 498, App | 761 | 7 | 0.8 | 2099 | 15 | US-10-128-714-3290 | Sequence 3290, Ap |
| 689 | 7 | 0.8 | 1041 | 12 | US-10-017-191A-498 | Sequence 498, App | 762 | 7 | 0.8 | 2405 | 15 | US-10-128-714-8290 | Sequence 8290, Ap |
| 690 | 7 | 0.8 | 1041 | 12 | US-10-143-028A-498 | Sequence 498, App | 763 | 7 | 0.8 | 2427 | 15 | US-10-369-493-6734 | Sequence 6734, Ap |
| 691 | 7 | 0.8 | 1041 | 12 | US-10-143-029A-498 | Sequence 498, App | 764 | 7 | 0.8 | 2630 | 12 | US-10-369-493-3693 | Sequence 3693, Ap |
| 692 | 7 | 0.8 | 1041 | 12 | US-10-145-089A-498 | Sequence 498, App | 765 | 7 | 0.8 | 3234 | 12 | US-10-093-463-1693 | Sequence 1693, App |
| 693 | 7 | 0.8 | 1041 | 12 | US-10-013-926A-498 | Sequence 498, App | 766 | 7 | 0.8 | 3390 | 12 | US-10-093-463-21101 | Sequence 21101, A |
| 694 | 7 | 0.8 | 1041 | 12 | US-10-145-017A-498 | Sequence 498, App | 767 | 7 | 0.8 | 3430 | 12 | US-10-369-493-3473 | Sequence 3473, Ap |
| 695 | 7 | 0.8 | 1041 | 12 | US-10-164-728A-498 | Sequence 498, App | 768 | 7 | 0.8 | 4999 | 10 | US-10-224-099A-3473 | Sequence 14, Appl |
| 696 | 7 | 0.8 | 1041 | 12 | US-10-165-067A-498 | Sequence 498, App | 769 | 7 | 0.8 | 19695 | 12 | US-09-976-059-14 | Sequence 3, Appl |
| 697 | 7 | 0.8 | 1041 | 12 | US-10-145-124A-498 | Sequence 498, App | 770 | 6 | 0.6 | 8 | 12 | US-10-084-846A-3 | Sequence 1592, Ap |
| 698 | 7 | 0.8 | 1041 | 12 | US-10-160-502A-498 | Sequence 498, App | 771 | 6 | 0.6 | 9 | 12 | US-10-351-641-1592 | Sequence 56, Appl |
| 699 | 7 | 0.8 | 1041 | 12 | US-10-165-247A-498 | Sequence 498, App | 772 | 6 | 0.6 | 9 | 12 | US-10-034-974-56 | Sequence 530, App |
| 700 | 7 | 0.8 | 1041 | 12 | US-09-978-194A-498 | Sequence 498, App | 773 | 6 | 0.6 | 9 | 12 | US-09-793-451-530 | Sequence 642, App |
| 701 | 7 | 0.8 | 1041 | 12 | US-09-978-681A-498 | Sequence 498, App | 774 | 6 | 0.6 | 9 | 12 | US-09-793-451-642 | Sequence 530, App |
| 702 | 7 | 0.8 | 1041 | 12 | US-09-999-829A-498 | Sequence 498, App | 775 | 6 | 0.6 | 9 | 12 | US-10-283-722-530 | Sequence 642, App |
| 703 | 7 | 0.8 | 1041 | 12 | US-10-013-922A-498 | Sequence 498, App | 776 | 6 | 0.6 | 9 | 12 | US-10-283-722-642 | Sequence 530, App |
| 704 | 7 | 0.8 | 1041 | 12 | US-10-017-086A-498 | Sequence 498, App | 777 | 6 | 0.6 | 9 | 12 | US-10-283-903-530 | Sequence 642, App |
| 705 | 7 | 0.8 | 1041 | 12 | US-10-145-087A-498 | Sequence 498, App | 778 | 6 | 0.6 | 10 | 7 | US-10-283-903-642 | Sequence 61, Appl |
| 706 | 7 | 0.8 | 1041 | 12 | US-10-164-829A-498 | Sequence 498, App | 779 | 6 | 0.6 | 10 | 11 | US-08-344-824-61 | Sequence 1963, Ap |
| 707 | 7 | 0.8 | 1041 | 12 | US-10-164-929A-498 | Sequence 498, App | 780 | 6 | 0.6 | 10 | 11 | US-09-572-404B-1963 | Sequence 1057, Ap |
| 708 | 7 | 0.8 | 1041 | 12 | US-09-978-299A-498 | Sequence 498, App | 781 | 6 | 0.6 | 10 | 12 | US-09-572-270A-1057 | Sequence 77, Appl |
| 709 | 7 | 0.8 | 1041 | 12 | US-09-978-544A-498 | Sequence 498, App | 782 | 6 | 0.6 | 10 | 12 | US-09-793-451-77 | Sequence 81, Appl |
| 710 | 7 | 0.8 | 1041 | 12 | US-09-978-665A-498 | Sequence 498, App | 783 | 6 | 0.6 | 10 | 12 | US-09-793-451-81 | Sequence 412, App |
| 711 | 7 | 0.8 | 1041 | 12 | US-09-978-802A-498 | Sequence 498, App | 784 | 6 | 0.6 | 10 | 12 | US-09-793-451-412 | Sequence 77, Appl |
| 712 | 7 | 0.8 | 1041 | 12 | US-10-013-924A-498 | Sequence 498, App | 785 | 6 | 0.6 | 10 | 12 | US-10-283-722-81 | Sequence 81, Appl |
| 713 | 7 | 0.8 | 1041 | 12 | US-10-020-445A-498 | Sequence 498, App | 786 | 6 | 0.6 | 10 | 12 | US-10-283-722-412 | Sequence 412, App |
| 714 | 7 | 0.8 | 1041 | 12 | US-10-017-084A-498 | Sequence 498, App | 787 | 6 | 0.6 | 10 | 12 | US-10-283-903-77 | Sequence 433, App |
| 715 | 7 | 0.8 | 1041 | 12 | US-10-017-085A-498 | Sequence 498, App | 788 | 6 | 0.6 | 10 | 12 | US-10-117-937-433 | Sequence 77, Appl |
| 716 | 7 | 0.8 | 1041 | 12 | US-10-143-916A-498 | Sequence 498, App | 789 | 6 | 0.6 | 10 | 12 | US-10-283-903-81 | Sequence 81, Appl |
| 717 | 7 | 0.8 | 1041 | 12 | US-10-013-926B-498 | Sequence 498, App | 790 | 6 | 0.6 | 10 | 12 | US-10-283-903-412 | Sequence 412, App |
| 718 | 7 | 0.8 | 1041 | 12 | US-10-013-918A-498 | Sequence 498, App | 791 | 6 | 0.6 | 10 | 15 | US-10-094-699-88 | Sequence 88, Appl |
| 719 | 7 | 0.8 | 1041 | 12 | US-10-013-923A-498 | Sequence 498, App | 792 | 6 | 0.6 | 10 | 15 | US-10-094-699-9 | Sequence 9, Appl |
| 720 | 7 | 0.8 | 1041 | 12 | US-10-013-925A-498 | Sequence 498, App | 793 | 6 | 0.6 | 10 | 15 | US-10-279-991-9 | Sequence 30, Appl |
| 721 | 7 | 0.8 | 1041 | 12 | US-10-013-927A-498 | Sequence 498, App | 794 | 6 | 0.6 | 12 | 12 | US-09-898-860-30 | Sequence 7, Appl |
| 722 | 7 | 0.8 | 1041 | 12 | US-10-013-928A-498 | Sequence 498, App | 795 | 6 | 0.6 | 12 | 12 | US-10-366-125-7 | Sequence 19, Appl |
| 723 | 7 | 0.8 | 1041 | 12 | US-10-162-522A-498 | Sequence 498, App | 796 | 6 | 0.6 | 13 | 14 | US-10-015-536-19 | Sequence 21, Appl |
| 724 | 7 | 0.8 | 1041 | 12 | US-10-407-352-28 | Sequence 28, Appl | 797 | 6 | 0.6 | 13 | 14 | US-10-015-536-21 | Sequence 40, Appl |
| 725 | 7 | 0.8 | 1041 | 12 | US-10-013-919A-498 | Sequence 498, App | 798 | 6 | 0.6 | 15 | 12 | US-10-034-974-40 | Sequence 6, Appl |
| 726 | 7 | 0.8 | 1041 | 12 | US-10-013-920A-498 | Sequence 498, App | 799 | 6 | 0.6 | 15 | 9 | US-10-279-991-6 | Sequence 14, Appl |
| 727 | 7 | 0.8 | 1041 | 15 | US-10-145-093A-498 | Sequence 3, Appl | 800 | 6 | 0.6 | 16 | 12 | US-09-770-940-14 | Sequence 14, Appl |
| 728 | 7 | 0.8 | 1041 | 15 | US-10-095-627-3 | Sequence 498, App | 801 | 6 | 0.6 | 16 | 12 | US-10-165-528-14 | Sequence 235, App |
| 729 | 7 | 0.8 | 1041 | 15 | US-10-017-081A-498 | Sequence 498, App | 802 | 6 | 0.6 | 17 | 12 | US-10-280-066-235 | Sequence 22, Appl |
| 730 | 7 | 0.8 | 1041 | 15 | US-10-167-749-498 | Sequence 498, App | 803 | 6 | 0.6 | 17 | 12 | US-10-425-328-22 | Sequence 22, Appl |
| 731 | 7 | 0.8 | 1041 | 15 | US-10-013-921A-498 | Sequence 498, App | 804 | 6 | 0.6 | 18 | 12 | US-10-145-206-130 | Sequence 130, App |
| 732 | 7 | 0.8 | 1041 | 15 | US-10-013-929A-498 | Sequence 498, App | 805 | 6 | 0.6 | 18 | 12 | US-10-408-166-454 | Sequence 454, App |
| 733 | 7 | 0.8 | 1041 | 15 | US-10-016-177A-498 | Sequence 498, App | 806 | 6 | 0.6 | 19 | 12 | US-10-408-166-453 | Sequence 453, App |
| 734 | 7 | 0.8 | 1041 | 15 | US-10-235-767-3 | Sequence 3, Appl | 807 | 6 | 0.6 | 20 | 11 | US-09-858-935B-89 | Sequence 89, Appl |
| 735 | 7 | 0.8 | 1041 | 15 | US-10-166-709A-498 | Sequence 498, App | 808 | 6 | 0.6 | 20 | 15 | US-10-408-166-452 | Sequence 452, App |
| 736 | 7 | 0.8 | 1059 | 11 | US-09-954-987B-187 | Sequence 187, App | 809 | 6 | 0.6 | 20 | 15 | US-10-044-708A-10 | Sequence 10, Appl |
| 737 | 7 | 0.8 | 1059 | 12 | US-10-407-952-30 | Sequence 30, Appl | 810 | 6 | 0.6 | 21 | 12 | US-10-408-166-451 | Sequence 451, App |
| 738 | 7 | 0.8 | 1109 | 10 | US-09-529-063-55 | Sequence 55, Appl | 811 | 6 | 0.6 | 22 | 12 | US-10-408-166-450 | Sequence 450, App |
| 739 | 7 | 0.8 | 1109 | 12 | US-10-414-378-55 | Sequence 55, Appl | 812 | 6 | 0.6 | 23 | 12 | US-10-408-166-449 | Sequence 449, App |
| 740 | 7 | 0.8 | 1140 | 12 | US-10-369-493-18194 | Sequence 18194, A | 813 | 6 | 0.6 | 24 | 9 | US-09-789-404-16 | Sequence 16, Appl |
| 741 | 7 | 0.8 | 1211 | 12 | US-10-452-024-107 | Sequence 107, App | 814 | 6 | 0.6 | 24 | 9 | US-09-875-494-1 | Sequence 1, Appl |
| 742 | 7 | 0.8 | 1247 | 10 | US-09-738-626-4751 | Sequence 824, App | 815 | 6 | 0.6 | 24 | 9 | US-09-875-494-2 | Sequence 2, Appl |
| 743 | 7 | 0.8 | 1266 | 12 | US-10-320-797-3352 | Sequence 4751, App | 816 | 6 | 0.6 | 24 | 12 | US-09-896-095-251 | Sequence 251, App |
| 744 | 7 | 0.8 | 1275 | 15 | US-10-128-714-3125 | Sequence 3352, App | 817 | 6 | 0.6 | 24 | 12 | US-10-273-973-9 | Sequence 9, Appl |
| 745 | 7 | 0.8 | 1280 | 12 | US-10-452-024-162 | Sequence 3125, App | 818 | 6 | 0.6 | 24 | 12 | US-10-408-166-448 | Sequence 448, App |

| | | | | | | | | | | | | | |
|-----|---|-----|----|----|---------------------|--------------------|-----|---|-----|----|----|---------------------|--------------------|
| 819 | 6 | 0.6 | 25 | 12 | US-10-408-166-447 | Sequence 447, App | 892 | 6 | 0.6 | 50 | 12 | US-10-408-166-374 | Sequence 374, App |
| 820 | 6 | 0.6 | 26 | 9 | US-09-864-761-45335 | Sequence 45335, A | 893 | 6 | 0.6 | 50 | 12 | US-10-408-166-422 | Sequence 422, App |
| 821 | 6 | 0.6 | 26 | 12 | US-10-408-166-446 | Sequence 446, App | 894 | 6 | 0.6 | 50 | 15 | US-10-106-698-8298 | Sequence 8298, App |
| 822 | 6 | 0.6 | 26 | 15 | US-10-012-542-383 | Sequence 383, App | 895 | 6 | 0.6 | 51 | 10 | US-09-989-903-34 | Sequence 34, Appl |
| 823 | 6 | 0.6 | 27 | 12 | US-10-408-166-445 | Sequence 445, App | 896 | 6 | 0.6 | 51 | 12 | US-10-408-166-373 | Sequence 373, Appl |
| 824 | 6 | 0.6 | 28 | 9 | US-09-770-940-15 | Sequence 15, Appl | 897 | 6 | 0.6 | 51 | 15 | US-10-068-564-34 | Sequence 34, Appl |
| 825 | 6 | 0.6 | 28 | 9 | US-09-864-761-34971 | Sequence 34971, A | 898 | 6 | 0.6 | 52 | 9 | US-09-864-761-42840 | Sequence 42840, A |
| 826 | 6 | 0.6 | 28 | 9 | US-09-864-761-41533 | Sequence 41533, A | 899 | 6 | 0.6 | 53 | 12 | US-10-105-232-370 | Sequence 370, App |
| 827 | 6 | 0.6 | 28 | 12 | US-10-165-528-15 | Sequence 15, Appl | 900 | 6 | 0.6 | 53 | 12 | US-10-189-437-357 | Sequence 357, App |
| 828 | 6 | 0.6 | 28 | 12 | US-10-408-166-444 | Sequence 444, App | 901 | 6 | 0.6 | 53 | 12 | US-10-408-166-189 | Sequence 189, App |
| 829 | 6 | 0.6 | 29 | 12 | US-10-029-386-31377 | Sequence 31377, A | 902 | 6 | 0.6 | 53 | 15 | US-10-106-698-5717 | Sequence 5717, App |
| 830 | 6 | 0.6 | 29 | 12 | US-10-408-166-443 | Sequence 443, App | 903 | 6 | 0.6 | 54 | 12 | US-09-864-408A-7516 | Sequence 7516, App |
| 831 | 6 | 0.6 | 30 | 9 | US-09-789-561-157 | Sequence 157, App | 904 | 6 | 0.6 | 54 | 12 | US-10-408-166-190 | Sequence 190, App |
| 832 | 6 | 0.6 | 30 | 10 | US-09-790-622-6 | Sequence 6, Appli | 905 | 6 | 0.6 | 55 | 9 | US-09-764-878-136 | Sequence 136, App |
| 833 | 6 | 0.6 | 30 | 12 | US-10-408-166-442 | Sequence 442, App | 906 | 6 | 0.6 | 55 | 9 | US-09-764-860-473 | Sequence 473, App |
| 834 | 6 | 0.6 | 30 | 15 | US-10-141-953-6 | Sequence 6, Appli | 907 | 6 | 0.6 | 55 | 12 | US-10-212-872-473 | Sequence 473, App |
| 835 | 6 | 0.6 | 31 | 9 | US-09-864-761-35950 | Sequence 35950, A | 908 | 6 | 0.6 | 55 | 12 | US-10-408-166-191 | Sequence 191, App |
| 836 | 6 | 0.6 | 31 | 12 | US-10-408-166-441 | Sequence 441, App | 909 | 6 | 0.6 | 55 | 15 | US-10-079-854-136 | Sequence 136, App |
| 837 | 6 | 0.6 | 32 | 12 | US-10-408-166-440 | Sequence 440, App | 910 | 6 | 0.6 | 55 | 15 | US-10-074-095-473 | Sequence 473, App |
| 838 | 6 | 0.6 | 33 | 9 | US-09-864-761-35444 | Sequence 35444, A | 911 | 6 | 0.6 | 56 | 12 | US-10-408-166-192 | Sequence 192, App |
| 839 | 6 | 0.6 | 33 | 12 | US-10-408-166-439 | Sequence 439, App | 912 | 6 | 0.6 | 57 | 12 | US-10-029-386-30158 | Sequence 30158, A |
| 840 | 6 | 0.6 | 33 | 14 | US-10-038-045-14 | Sequence 14, Appl | 913 | 6 | 0.6 | 57 | 12 | US-10-408-166-193 | Sequence 193, App |
| 841 | 6 | 0.6 | 34 | 12 | US-10-408-166-438 | Sequence 438, App | 914 | 6 | 0.6 | 58 | 10 | US-09-948-080-22 | Sequence 22, Appl |
| 842 | 6 | 0.6 | 35 | 12 | US-10-408-166-437 | Sequence 437, App | 915 | 6 | 0.6 | 58 | 12 | US-10-156-761-8226 | Sequence 8226, App |
| 843 | 6 | 0.6 | 35 | 14 | US-10-029-217A-23 | Sequence 23, Appl | 916 | 6 | 0.6 | 58 | 15 | US-10-156-761-8226 | Sequence 8226, App |
| 844 | 6 | 0.6 | 36 | 9 | US-09-864-761-44661 | Sequence 44661, A | 917 | 6 | 0.6 | 59 | 11 | US-09-892-877-339 | Sequence 339, App |
| 845 | 6 | 0.6 | 36 | 12 | US-10-408-166-436 | Sequence 436, App | 918 | 6 | 0.6 | 59 | 11 | US-09-948-783-352 | Sequence 352, App |
| 846 | 6 | 0.6 | 37 | 9 | US-09-864-761-43447 | Sequence 43447, A | 919 | 6 | 0.6 | 59 | 12 | US-10-408-166-195 | Sequence 195, App |
| 847 | 6 | 0.6 | 37 | 11 | US-09-776-724A-88 | Sequence 88, Appl | 920 | 6 | 0.6 | 60 | 12 | US-10-408-166-196 | Sequence 196, App |
| 848 | 6 | 0.6 | 37 | 12 | US-10-029-386-33604 | Sequence 33604, A | 921 | 6 | 0.6 | 60 | 15 | US-10-083-357-817 | Sequence 817, App |
| 849 | 6 | 0.6 | 37 | 12 | US-10-408-166-435 | Sequence 435, App | 922 | 6 | 0.6 | 61 | 9 | US-09-764-869-1160 | Sequence 1160, App |
| 850 | 6 | 0.6 | 38 | 10 | US-09-925-300-1127 | Sequence 1127, App | 923 | 6 | 0.6 | 61 | 12 | US-10-227-577-1160 | Sequence 1160, App |
| 851 | 6 | 0.6 | 38 | 12 | US-10-408-166-434 | Sequence 434, App | 924 | 6 | 0.6 | 61 | 12 | US-10-408-166-197 | Sequence 197, App |
| 852 | 6 | 0.6 | 39 | 12 | US-10-408-166-385 | Sequence 385, App | 925 | 6 | 0.6 | 61 | 15 | US-10-091-504-1160 | Sequence 1160, App |
| 853 | 6 | 0.6 | 39 | 12 | US-10-408-166-433 | Sequence 433, App | 926 | 6 | 0.6 | 62 | 11 | US-09-764-891-3703 | Sequence 3703, App |
| 854 | 6 | 0.6 | 40 | 12 | US-10-408-166-384 | Sequence 384, App | 927 | 6 | 0.6 | 62 | 12 | US-09-864-408A-5028 | Sequence 5028, App |
| 855 | 6 | 0.6 | 40 | 12 | US-10-408-166-432 | Sequence 432, App | 928 | 6 | 0.6 | 62 | 12 | US-10-408-166-198 | Sequence 198, App |
| 856 | 6 | 0.6 | 41 | 12 | US-10-408-166-383 | Sequence 383, App | 929 | 6 | 0.6 | 63 | 10 | US-09-738-626-3628 | Sequence 3628, App |
| 857 | 6 | 0.6 | 41 | 12 | US-10-408-166-431 | Sequence 431, App | 930 | 6 | 0.6 | 63 | 12 | US-10-408-166-199 | Sequence 199, App |
| 858 | 6 | 0.6 | 42 | 12 | US-10-408-166-382 | Sequence 382, App | 931 | 6 | 0.6 | 64 | 12 | US-09-933-787-570 | Sequence 570, App |
| 859 | 6 | 0.6 | 42 | 12 | US-10-408-166-430 | Sequence 430, App | 932 | 6 | 0.6 | 64 | 15 | US-10-408-166-200 | Sequence 200, App |
| 860 | 6 | 0.6 | 43 | 9 | US-09-864-761-43508 | Sequence 43508, A | 933 | 6 | 0.6 | 64 | 15 | US-10-023-282-570 | Sequence 570, App |
| 861 | 6 | 0.6 | 43 | 12 | US-10-321-857-108 | Sequence 108, App | 934 | 6 | 0.6 | 65 | 12 | US-10-029-386-33591 | Sequence 33591, A |
| 862 | 6 | 0.6 | 43 | 12 | US-10-318-675-108 | Sequence 108, App | 935 | 6 | 0.6 | 65 | 12 | US-10-408-166-201 | Sequence 201, App |
| 863 | 6 | 0.6 | 43 | 12 | US-10-408-166-381 | Sequence 381, App | 936 | 6 | 0.6 | 66 | 12 | US-09-864-408A-5204 | Sequence 5204, App |
| 864 | 6 | 0.6 | 44 | 12 | US-10-408-166-429 | Sequence 429, App | 937 | 6 | 0.6 | 66 | 12 | US-10-408-166-202 | Sequence 202, App |
| 865 | 6 | 0.6 | 44 | 11 | US-09-895-298-126 | Sequence 126, App | 938 | 6 | 0.6 | 67 | 12 | US-09-864-408A-186 | Sequence 186, App |
| 866 | 6 | 0.6 | 44 | 12 | US-10-195-730-339 | Sequence 339, App | 939 | 6 | 0.6 | 67 | 12 | US-10-408-166-203 | Sequence 203, App |
| 867 | 6 | 0.6 | 44 | 12 | US-10-408-166-428 | Sequence 380, App | 940 | 6 | 0.6 | 68 | 11 | US-09-823-187-47 | Sequence 47, Appl |
| 868 | 6 | 0.6 | 44 | 12 | US-10-408-166-428 | Sequence 428, App | 941 | 6 | 0.6 | 68 | 12 | US-10-408-166-204 | Sequence 204, App |
| 869 | 6 | 0.6 | 45 | 9 | US-09-864-761-39649 | Sequence 39649, A | 942 | 6 | 0.6 | 69 | 9 | US-09-864-761-4802 | Sequence 4802, A |
| 870 | 6 | 0.6 | 45 | 9 | US-09-864-761-48064 | Sequence 48064, A | 943 | 6 | 0.6 | 69 | 10 | US-09-738-626-5569 | Sequence 5569, App |
| 871 | 6 | 0.6 | 45 | 12 | US-10-408-166-379 | Sequence 379, App | 944 | 6 | 0.6 | 69 | 12 | US-10-029-386-27796 | Sequence 27796, A |
| 872 | 6 | 0.6 | 45 | 12 | US-10-408-166-427 | Sequence 427, App | 945 | 6 | 0.6 | 69 | 12 | US-10-408-166-205 | Sequence 205, App |
| 873 | 6 | 0.6 | 46 | 9 | US-09-205-658-83 | Sequence 83, Appl | 946 | 6 | 0.6 | 70 | 11 | US-09-764-891-3021 | Sequence 3021, App |
| 874 | 6 | 0.6 | 46 | 9 | US-09-844-353A-83 | Sequence 83, Appl | 947 | 6 | 0.6 | 70 | 12 | US-10-408-166-206 | Sequence 206, App |
| 875 | 6 | 0.6 | 46 | 9 | US-09-864-761-35341 | Sequence 35341, A | 948 | 6 | 0.6 | 71 | 10 | US-09-943-123-15 | Sequence 15, Appl |
| 876 | 6 | 0.6 | 46 | 12 | US-09-963-693-83 | Sequence 83, Appl | 949 | 6 | 0.6 | 71 | 12 | US-10-408-166-207 | Sequence 207, App |
| 877 | 6 | 0.6 | 46 | 12 | US-10-408-166-378 | Sequence 378, App | 950 | 6 | 0.6 | 71 | 15 | US-10-050-704-112 | Sequence 112, App |
| 878 | 6 | 0.6 | 46 | 12 | US-10-408-166-426 | Sequence 426, App | 951 | 6 | 0.6 | 72 | 11 | US-09-764-891-2973 | Sequence 2973, App |
| 879 | 6 | 0.6 | 47 | 12 | US-10-029-386-31159 | Sequence 31159, A | 952 | 6 | 0.6 | 72 | 12 | US-10-266-829-103 | Sequence 103, App |
| 880 | 6 | 0.6 | 47 | 12 | US-10-408-166-377 | Sequence 377, App | 953 | 6 | 0.6 | 72 | 12 | US-10-408-166-208 | Sequence 208, App |
| 881 | 6 | 0.6 | 47 | 12 | US-10-408-166-425 | Sequence 425, App | 954 | 6 | 0.6 | 73 | 12 | US-10-408-166-209 | Sequence 209, App |
| 882 | 6 | 0.6 | 48 | 12 | US-09-933-767-358 | Sequence 358, App | 955 | 6 | 0.6 | 74 | 11 | US-09-468-147-190 | Sequence 190, App |
| 883 | 6 | 0.6 | 48 | 12 | US-10-408-166-376 | Sequence 376, App | 956 | 6 | 0.6 | 74 | 11 | US-09-468-147-197 | Sequence 197, App |
| 884 | 6 | 0.6 | 48 | 12 | US-10-408-166-424 | Sequence 424, App | 957 | 6 | 0.6 | 74 | 12 | US-10-029-386-28264 | Sequence 28264, A |
| 885 | 6 | 0.6 | 48 | 15 | US-10-023-282-358 | Sequence 358, App | 958 | 6 | 0.6 | 74 | 12 | US-09-864-408A-6794 | Sequence 6794, App |
| 886 | 6 | 0.6 | 49 | 9 | US-09-789-561-127 | Sequence 127, App | 959 | 6 | 0.6 | 74 | 14 | US-10-408-166-210 | Sequence 210, App |
| 887 | 6 | 0.6 | 49 | 12 | US-09-833-245-2102 | Sequence 2102, App | 960 | 6 | 0.6 | 74 | 14 | US-10-078-929-68 | Sequence 68, Appl |
| 888 | 6 | 0.6 | 49 | 12 | US-10-408-166-375 | Sequence 375, App | 961 | 6 | 0.6 | 75 | 9 | US-09-864-761-45644 | Sequence 45644, A |
| 889 | 6 | 0.6 | 49 | 12 | US-10-408-166-423 | Sequence 423, App | 962 | 6 | 0.6 | 75 | 12 | US-10-408-166-211 | Sequence 211, App |
| 890 | 6 | 0.6 | 50 | 10 | US-09-911-826A-15 | Sequence 15, Appl | 963 | 6 | 0.6 | 76 | 9 | US-09-864-761-44615 | Sequence 44615, A |
| 891 | 6 | 0.6 | 50 | 12 | US-10-190-435-197 | Sequence 197, App | 964 | 6 | 0.6 | 76 | 10 | US-09-963-959-10 | Sequence 10, Appl |

| | | | | | | | | | | | |
|---|------|---|-----|----|----|---------------------|-------------------|----|-----|--|-----|
| 6 | 965 | 6 | 0.6 | 76 | 12 | US-10-408-166-212 | Sequence 212, App | Db | 61 | NHITAKSVDTVNOQFSLTQTGTIAISATKLEKPLQKHSTNKLAKGLDVENIDRKLGKA | 120 |
| 6 | 966 | 6 | 0.6 | 76 | 15 | US-10-156-761-9828 | Sequence 9828, Ap | Qy | 121 | SNVLSTLSSFLGTALAGIELDSLKKGDAAPDALAKASIDLINELIIGNLSQSTOTIEAFS | 180 |
| 6 | 967 | 6 | 0.6 | 77 | 9 | US-09-939-980-511 | Sequence 511, App | Qy | 121 | SNVLSTLSSFLGTALAGIELDSLKKGDAAPDALAKASIDLINELIIGNLSQSTOTIEAFS | 180 |
| 6 | 968 | 6 | 0.6 | 77 | 11 | US-09-963-959-5 | Sequence 5, Appl1 | Db | 121 | SNVLSTLSSFLGTALAGIELDSLKKGDAAPDALAKASIDLINELIIGNLSQSTOTIEAFS | 180 |
| 6 | 969 | 6 | 0.6 | 77 | 11 | US-09-764-891-5129 | Sequence 5129, Ap | Qy | 181 | SOLAKLGSSTTSQAKGFSNIGNKQLNLFNSKTNLGLIITGLLSGISAGFALADKNASTGK | 240 |
| 6 | 970 | 6 | 0.6 | 77 | 12 | US-10-408-166-213 | Sequence 213, App | Qy | 181 | SOLAKLGSSTTSQAKGFSNIGNKQLNLFNSKTNLGLIITGLLSGISAGFALADKNASTGK | 240 |
| 6 | 971 | 6 | 0.6 | 78 | 9 | US-09-764-869-800 | Sequence 800, App | Db | 241 | KVAAGFELSNQVIGNVTKAISYVLAQKRVAAAGLSTTGAVAAALITSSIMLAISPLAFMNA | 300 |
| 6 | 972 | 6 | 0.6 | 78 | 12 | US-10-227-577-800 | Sequence 800, App | Qy | 241 | KVAAGFELSNQVIGNVTKAISYVLAQKRVAAAGLSTTGAVAAALITSSIMLAISPLAFMNA | 300 |
| 6 | 973 | 6 | 0.6 | 78 | 12 | US-10-408-166-214 | Sequence 214, App | Db | 301 | DKFNHANALDEFAKQFRKFGYDGDHLLAEYQGVGTTEASLTTTSTALGAVSAGVSA | 360 |
| 6 | 974 | 6 | 0.6 | 78 | 15 | US-10-091-504-800 | Sequence 800, App | Qy | 301 | DKFNHANALDEFAKQFRKFGYDGDHLLAEYQGVGTTEASLTTTSTALGAVSAGVSA | 360 |
| 6 | 975 | 6 | 0.6 | 79 | 9 | US-09-764-887-216 | Sequence 216, App | Qy | 361 | GSVGTPTALIVAGVTGLISGILEASQKAMPESVANRLQKILBEWKQNGQNYFDKGYD | 420 |
| 6 | 976 | 6 | 0.6 | 79 | 10 | US-09-738-626-5321 | Sequence 5321, Ap | Qy | 361 | GSVGTPTALIVAGVTGLISGILEASQKAMPESVANRLQKILBEWKQNGQNYFDKGYD | 420 |
| 6 | 977 | 6 | 0.6 | 79 | 12 | US-10-408-166-215 | Sequence 215, App | Db | 421 | SRYAAYLANNLKPLSELNKELEASERVIAITQORWNNNIGELAGITKLGRIKSGKAYADA | 480 |
| 6 | 978 | 6 | 0.6 | 79 | 15 | US-10-073-961-216 | Sequence 216, App | Qy | 421 | SRYAAYLANNLKPLSELNKELEASERVIAITQORWNNNIGELAGITKLGRIKSGKAYADA | 480 |
| 6 | 979 | 6 | 0.6 | 80 | 12 | US-10-408-166-216 | Sequence 216, App | Qy | 481 | FEDGKVEAGSNITLDAKTGIIDISNSGKKTQALHTSPLLTGTAGTSREBLTNGKYSYI | 540 |
| 6 | 980 | 6 | 0.6 | 81 | 11 | US-09-992-600A-38 | Sequence 38, Appl | Db | 481 | FEDGKVEAGSNITLDAKTGIIDISNSGKKTQALHTSPLLTGTAGTSREBLTNGKYSYI | 540 |
| 6 | 981 | 6 | 0.6 | 81 | 11 | US-09-924-340-38 | Sequence 38, Appl | Qy | 541 | NKLKFRVKWQVTDGEASSKLDFESKVIQORVAETEGTDEIGLI VNAKAGNDDIIFVGQGM | 600 |
| 6 | 982 | 6 | 0.6 | 81 | 12 | US-09-992-095B-38 | Sequence 38, Appl | Qy | 541 | NKLKFRVKWQVTDGEASSKLDFESKVIQORVAETEGTDEIGLI VNAKAGNDDIIFVGQGM | 600 |
| 6 | 983 | 6 | 0.6 | 81 | 12 | US-10-154-678-38 | Sequence 38, Appl | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 984 | 6 | 0.6 | 81 | 12 | US-09-999-570-38 | Sequence 38, Appl | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 985 | 6 | 0.6 | 81 | 12 | US-10-289-762-1167 | Sequence 1167, Ap | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 986 | 6 | 0.6 | 81 | 12 | US-10-408-166-217 | Sequence 217, Ap | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 987 | 6 | 0.6 | 81 | 15 | US-10-000-489-38 | Sequence 38, Appl | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 988 | 6 | 0.6 | 81 | 15 | US-10-000-986-38 | Sequence 38, Appl | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 989 | 6 | 0.6 | 82 | 12 | US-10-408-166-218 | Sequence 218, App | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 990 | 6 | 0.6 | 82 | 15 | US-10-301-084-34 | Sequence 34, Appl | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 991 | 6 | 0.6 | 83 | 12 | US-09-864-408A-5086 | Sequence 5086, Ap | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 992 | 6 | 0.6 | 83 | 12 | US-10-408-166-219 | Sequence 219, App | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 993 | 6 | 0.6 | 84 | 12 | US-09-833-245-1908 | Sequence 1908, Ap | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 994 | 6 | 0.6 | 84 | 12 | US-09-833-245-1909 | Sequence 1909, Ap | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 995 | 6 | 0.6 | 84 | 12 | US-10-408-166-220 | Sequence 220, App | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 996 | 6 | 0.6 | 85 | 9 | US-09-925-297-462 | Sequence 462, App | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 997 | 6 | 0.6 | 85 | 12 | US-10-029-386-30397 | Sequence 30397, A | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 998 | 6 | 0.6 | 85 | 12 | US-10-408-166-221 | Sequence 221, App | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 999 | 6 | 0.6 | 86 | 12 | US-10-029-386-31914 | Sequence 31914, A | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |
| 6 | 1000 | 6 | 0.6 | 86 | 12 | US-10-316-194-48 | Sequence 48, Appl | Qy | 601 | NIDGGDGHDRVYFSKDGOGFGNITVDGTSATGAGSYTVNRKVGARDIYHEVVVQRETQKVGK | 660 |

ALIGNMENTS

| QY | 721 | GAGDDRLFCGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN | 780 |
|----|-----|--|-----|
| QY | 721 | GAGDDRLFCGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN | 780 |
| Db | 721 | GAGDDRLFCGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN | 780 |
| QY | 781 | DKLAFADANISDIMIERTKEGIIIVKRNDSHSINIPRWYITISNLQYOSNKTDKHIEQLI | 840 |
| Db | 781 | DKLAFADANISDIMIERTKEGIIIVKRNDSHSINIPRWYITISNLQYOSNKTDKHIEQLI | 840 |
| QY | 841 | GKDGSVITSDDTKILODKOGTVITTSOELKKLADENKSOKLSASDIASSINKLVGSNAL | 900 |
| Db | 841 | GKDGSVITSDDTKILODKOGTVITTSOELKKLADENKSOKLSASDIASSINKLVGSNAL | 900 |
| QY | 901 | FGTANSVSSNALQPITQPTQGILAPSV | 927 |
| Db | 901 | FGTANSVSSNALQPITQPTQGILAPSV | 927 |

RESULT 2

US-09-884-696-13

; Sequence 13, Application US/09884696

; Publication No. US20030035809A1

; GENERAL INFORMATION:

; APPLICANT: GEORGE, LISLE W

; APPLICANT: ANGELOS, JOHN A

; APPLICANT: HESS, JOHN F

; TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES

; TITLE OF INVENTION: AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA

; FILE OF INVENTION: BOVIS INFECTIONS

; FILE REFERENCE: 481.06

; CURRENT APPLICATION NUMBER: US/09/884,696

; CURRENT FILING DATE: 2001-06-19

US-09-884-696-2

; Sequence 2, Application US/09884696

; Publication No. US20030035809A1

; GENERAL INFORMATION:

; APPLICANT: GEORGE, LISLE W

; APPLICANT: ANGELOS, JOHN A

; APPLICANT: HESS, JOHN F

; TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES

; TITLE OF INVENTION: AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA

; FILE OF INVENTION: BOVIS INFECTIONS

; FILE REFERENCE: 481.06

; CURRENT APPLICATION NUMBER: US/09/884,696

; CURRENT FILING DATE: 2001-06-19

NUMBER OF SEQ ID NOS: 41

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 927

TYPE: PRT

ORGANISM: Moraxella bovis

US-09-884-696-2

Query Match 89.1%; Score 826; DB 11; Length 927;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 926; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSNINVIKSNIQAGLNSTKSGLNLYLAIPKDYDPQKGGTNDFTKAADELGIARLAEPP 60

Db 1 MSNINVIKSNIQAGLNSTKSGLNLYLAIPKDYDPQKGGTNDFTKAADELGIARLAEPP 60

QY 61 NHTETAKSVDTWNOFLSTQGTGIAISATKLEKFKQKSTNKLAKGLDSVENIDRLGLKA 120

[illegible]

; APPLICANT: GEORGE, LISLE W
; APPLICANT: ANGELOS, JOHN A
; APPLICANT: HESS, JOHN F
; TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
; TITLE OF INVENTION: AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
; TITLE OF INVENTION: BOVIS INFECTIONS
; FILE REFERENCE: 481.06
; CURRENT APPLICATION NUMBER: US/09/884,696
; CURRENT FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 3
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Pasteurella haemolytica
US-09-884-696-3

Query Match 1.1%; Score 10; DB 11; Length 953;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 LAQRYAAGLS 274
| | | | | | | |
Db 283 LAQRYAAGLS 292

RESULT 8
US-10-369-493-20166
; Sequence 20166, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20166
; LENGTH: 1017
; TYPE: PRT
; ORGANISM: No. US20030233675A1toc punctiforme
; NAME/KEY: unsure
; LOCATION: (1)-(1017)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-20166

Query Match 1.1%; Score 10; DB 12; Length 1017;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 718 LDGGAGDDRL 727
| | | | | | | |
Db 389 LDGGAGDDRL 398

RESULT 9
US-10-156-761-12748
; Sequence 12748, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12748
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12748

Query Match 1.0%; Score 9; DB 15; Length 259;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 276 TGAAVALIT 284
| | | | | | | |
Db 164 TGAAVALIT 172

RESULT 10
US-10-156-761-12788
; Sequence 12788, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12788
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12788

Query Match 0.9%; Score 8; DB 15; Length 307;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 353 AGVSAAAV 360
| | | | | | | |
Db 109 AGVSAAAV 116

RESULT 11
US-10-156-761-14641
; Sequence 14641, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14641
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14641

Query Match 0.9%; Score 8; DB 15; Length 408;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 357 AAAGSAV 364
|||||
DB 298 AAAGSAV 305

RESULT 12
US-10-369-493-17978
; Sequence 17978, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17978
; LENGTH: 424
; TYPE: PRT
; ORGANISM: SPHINGOMONAS
; NAME/KEY: unsure
; FEATURE:
; LOCATION: (1)-(424)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-17978

Query Match 0.9%; Score 8; DB 12; Length 424;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 AAPDALAK 156
|||||
DB 247 AAPDALAK 254

RESULT 13
US-10-369-493-18633
; Sequence 18633, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18633
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Halobacterium sp. NRC-1
US-10-369-493-18633

Query Match 0.9%; Score 8; DB 12; Length 477;
Best Local Similarity 100.0%; Pred. No. 111e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 AAVGSAVG 365
|||||
DB 15 AAVGSAVG 22

RESULT 14
US-10-369-493-16555
; Sequence 16555, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16555
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
; NAME/KEY: unsure
; FEATURE:
; LOCATION: (1)-(489)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-16555

Query Match 0.9%; Score 8; DB 12; Length 489;
Best Local Similarity 100.0%; Pred. No. 1.11e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 AAVGSAVG 365
|||||
DB 15 AAVGSAVG 22

RESULT 15
US-10-369-493-17426
; Sequence 17426, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17426
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Bacillus halodurans
US-10-369-493-17426

Query Match 0.9%; Score 8; DB 12; Length 505;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 358 AAVGSAGV 365
Db 14 AAVGSAGV 21

RESULT 16
US-10-156-761-11340
; Sequence 11340, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156.761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11340
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11340

Query Match 0.9%; Score 8; DB 15; Length 527;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 347 ALGAVSAG 354
Db 75 ALGAVSAG 82

RESULT 17
US-10-369-493-1471
; Sequence 1471, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 1471
; LENGTH: 773
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1471

Query Match 0.9%; Score 8; DB 12; Length 773;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 KAADELGI 53
Db 667 KAADELGI 674

RESULT 18
US-10-369-493-10178
; Sequence 10178, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10178
; LENGTH: 4327
; TYPE: PRT
; ORGANISM: magnetite-containing magnetic coccus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(4327)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-10178

Query Match 0.9%; Score 8; DB 12; Length 4327;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 720 GGAGDDRL 727
Db 3976 GGAGDDRL 3983

RESULT 19
US-10-414-692-80
; Sequence 80, Application US/10414692
; Publication No. US20030228607A1
; GENERAL INFORMATION:
; APPLICANT: X-ceptor Therapeutics, Inc.
; TITLE OF INVENTION: Screening method and modulators having an improved therapeutic
; FILE REFERENCE: 8012-002-US
; CURRENT APPLICATION NUMBER: US/10/414,692
; CURRENT FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 60/372,650
; PRIOR FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 80
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic co-activator peptide

US-10-414-692-80

Query Match 0.8%; Score 7; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 744 LLDGGG 750
|||||
Db 12 LLDGGG 18

RESULT 20

US-08-424-550B-519
; Sequence 519, Application US/08424550B
; Publication No. US20020119447A1

GENERAL INFORMATION:

; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHRHOFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR

; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE

; NUMBER OF SEQUENCES: 716

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D

; STREET: 100 ABBOTT PARK ROAD

; CITY: ABBOTT PARK

; STATE: IL

; COUNTRY: USA

; ZIP: 60064-3500

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/424,550B

; FILING DATE:

; CLASSIFICATION: 435435

; ATTORNEY/AGENT INFORMATION:

; NAME: FOREMSKI, PRISCILLA E.

; REGISTRATION NUMBER: 33,207

; REFERENCE/DOCKET NUMBER: 5527.PC.01

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 708-937-6365

; TELEFAX: 708-938-2623

; INFORMATION FOR SEQ ID NO: 519:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 44 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-424-550B-519

Query Match 0.8%; Score 7; DB 8; Length 44;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 898 MALFGTA 904
|||||
Db 38 MALFGTA 44

RESULT 21

US-10-264-049-3945

; Sequence 3945, Application US/10264049

; Publication No. US20040005579A1

GENERAL INFORMATION:

; APPLICANT: Birse et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PA133PI

; CURRENT APPLICATION NUMBER: US/10/264,049

; CURRENT FILING DATE: 2002-10-04

; PRIOR APPLICATION NUMBER: PCT/US01/18569

; PRIOR FILING DATE: 2001-06-07

; PRIOR APPLICATION NUMBER: US 60/209,467

; PRIOR FILING DATE: 2000-06-07

; NUMBER OF SEQ ID NOS: 4360

; SOFTWARE: PatentIn Ver. 3.1

; SEQ ID NO 3945

; LENGTH: 49

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-264-049-3945

Query Match .0.8%; Score 7; DB 12; Length 49;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 139 ELDSLIIK 145
|||||
Db 17 ELDSLIIK 23

RESULT 22

US-10-185-050-78

; Sequence 78, Application US/10185050

; Publication No. US2003007577A1

; GENERAL INFORMATION:

; APPLICANT: Pirozzi, Gregorio

; Kay, Brian K.

; Fowlkes, Dana M.

; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL

; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME

; NUMBER OF SEQUENCES: 233

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PENNIE & EDMONDS LLP

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/185,050

; FILING DATE: 28-Jun-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/826,516

; FILING DATE: 03-Apr-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: MISROCK, S. LESLIE

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 1101-208-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 896-8864/9741

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 78:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 52 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 78:

US-10-185-050-78

Query Match 0.8%; Score 7; DB 15; Length 52;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 243 AAGFELS 249
|||||||
Db 46 AAGFELS 52

RESULT 23

US-09-917-340-9
; Sequence 9, Application US/09917340
; Patent No. US2002090369A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Christopher J.
; APPLICANT: McNulty, Jonathan F.
; APPLICANT: Reid, Ted W.
; TITLE OF INVENTION: Transplant Media
; FILE REFERENCE: TPLANT-06468
; CURRENT APPLICATION NUMBER: US/09/917,340
; CURRENT FILING DATE: 2001-07-29
; PRIOR APPLICATION NUMBER: 60/221,632
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/249,602
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/290,932
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Bombyx mori

US-09-917-340-9

Query Match 0.8%; Score 7; DB 9; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 347 ALGAVSA 353
|||||||
Db 16 ALGAVSA 22

RESULT 24

US-09-739-907-83
; Sequence 83, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/739,907
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 83
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: SITE
; LOCATION: (89)
; OTHER INFORMATION: Xaa equals stop translation
US-09-739-907-83

Query Match 0.8%; Score 7; DB 9; Length 89;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 357 AAAGVSA 363
|||||||
Db 52 AAAGVSA 58

RESULT 25

US-09-938-671-83
; Sequence 83, Application US/09938671
; Publication No. US20040002066A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/938,671
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 83
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (89)
; OTHER INFORMATION: Xaa equals stop translation
US-09-938-671-83

Query Match 0.8%; Score 7; DB 12; Length 89;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 357 AAAGVSA 363
|||||||
Db 52 AAAGVSA 58

RESULT 26

US-09-864-408A-5506
; Sequence 5506, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5506
; LENGTH: 90
; TYPE: PRT

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-09-864-408A-5506

Query Match          0.8%; Score 7; DB 12; Length 90;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 445 RVIAITQ 451
Db 31 RVIAITQ 37

RESULT 27
US-10-108-260A-3335
; Sequence 3335, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3335
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3335

Query Match          0.8%; Score 7; DB 12; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 371 LVAGVTG 377
Db 64 LVAGVTG 70

RESULT 28
US-09-924-358-20
; Sequence 20, Application US/09924358
; Patent No. US20020107376A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 26199, 33530, 33949, 47148, 50226, AND
; TITLE OF INVENTION: 58764,
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 38155-20034.00
; CURRENT APPLICATION NUMBER: US/09/924,358
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/229,300
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid consensus sequence
US-10-410-764-20

Query Match          0.8%; Score 7; DB 12; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 761 DVYIFRK 767
Db 92 DVYIFRK 98

RESULT 30
US-09-815-242-11108
; Sequence 11108, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
US-10-410-764-20

Query Match          0.8%; Score 7; DB 12; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 761 DVYIFRK 767
Db 92 DVYIFRK 98

RESULT 30
US-09-815-242-11108
; Sequence 11108, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
```

APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 11108
LENGTH: 103
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-815-242-11108

Query Match 0.8%; Score 7; DB 9; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 481 FEDGKV 487
Db 87 FEDGKV 93

RESULT 31
US-09-741-669-297
Sequence 297, Application US/09741669
Patent No. US2002002718A1
GENERAL INFORMATION:
APPLICANT: Forsyth, R. Alllyn
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
TITLE OF INVENTION: Genes identified as required for
TITLE OF INVENTION: Proliferation of E. coli
FILE REFERENCE: ELITRA.009A
CURRENT APPLICATION NUMBER: US/09/741,669
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 60/173005
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 481
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 297
LENGTH: 104
TYPE: PRT
ORGANISM: Escherichia coli
US-09-741-669-297

Query Match 0.8%; Score 7; DB 9; Length 104;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 481 FEDGKV 487
Db 87 FEDGKV 93

RESULT 32
US-09-912-020-321
Sequence 321, Application US/09912020
Patent No. US2002004592A1
GENERAL INFORMATION:
APPLICANT: Zyskind, Judith
APPLICANT: Ohlsen, Kari L.
APPLICANT: Trawick, John
APPLICANT: Forsyth, R. Alllyn
APPLICANT: Froelich, Jamie M.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
TITLE OF INVENTION: ESCHERICHIA COLI
FILE REFERENCE: ELITRA.001DV1
CURRENT APPLICATION NUMBER: US/09/912,020
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: 09/492,709
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: 60/117,405
PRIOR FILING DATE: 1999-01-27
NUMBER OF SEQ ID NOS: 485
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 321Z
LENGTH: 104
TYPE: PRT
ORGANISM: E. Coli
US-09-912-020-321

Query Match 0.8%; Score 7; DB 9; Length 104;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 481 FEDGKV 487
Db 87 FEDGKV 93

RESULT 33
US-09-815-242-10348
Sequence 10348, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10348
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10348

Query Match          0.8%; Score 7; DB 9; Length 104;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      481 FEDGKKV 487
Db      87 FEDGKKV 93

RESULT 34
US-09-815-242-14107
; Sequence 14107, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14107
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(104)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-14107

Query Match          0.8%; Score 7; DB 9; Length 104;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      481 FEDGKKV 487
Db      87 FEDGKKV 93

RESULT 35
US-10-287-274-323
; Sequence 323, Application US/10287274
; Publication No. US20030181409A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERETO
; FILE REFERENCE: ELITRA.008DV1
; CURRENT APPLICATION NUMBER: US/10/287,274
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: US 09/711164
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 323
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-287-274-323

Query Match          0.8%; Score 7; DB 12; Length 104;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      481 FEDGKKV 487
Db      87 FEDGKKV 93

RESULT 36
US-09-989-919-93
; Sequence 93, Application US/09989919
; Patent No. US20020164344A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Pluta, Jason
; APPLICANT: Ghosh, Malavika
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chonghua
; TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and Pro
; FILE REFERENCE: DEX-0289
; CURRENT APPLICATION NUMBER: US/09/989,919
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,505
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 93
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-919-93

Query Match          0.8%; Score 7; DB 10; Length 109;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      559 SSKLDFS 565
Db      25 SSKLDFS 31

RESULT 37
US-09-864-761-40104
; Sequence 40104, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
```


; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40104
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004132.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.8
; OTHER INFORMATION: EST HUMAN HIT: AUI29622.1, EVALUE 5.00e-42
; OTHER INFORMATION: SWISSPROT HIT: P35680, EVALUE 5.00e-43
US-09-864-761-40104

Query Match 0.8%; Score 7; DB 9; Length 114;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 735 RLSGDEG 741

Db 73 RLSGDEG 79

RESULT 38

US-10-264-049-3972
; Sequence 3972, Application US/10264049
; Publication No. US20040005579A1

; GENERAL INFORMATION:

; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAI33PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3972
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (98)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (106)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-3972

Query Match 0.8%; Score 7; DB 12; Length 115;

Best Local Similarity 100.0%; Pred. No. 2.6e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 368 IALLVAG 374

Db 3 IALLVAG 9

RESULT 39

US-10-094-749-2223

; Sequence 2223, Application US/10094749

; Publication No. US20030219741A1

; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO

; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA

; FILE REFERENCE: 084335/0160

; CURRENT APPLICATION NUMBER: US/10/094,749

; CURRENT FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: 60/350,435

; PRIOR FILING DATE: 2002-01-24

; PRIOR APPLICATION NUMBER: JP 2001-328381

; PRIOR FILING DATE: 2001-09-14

; NUMBER OF SEQ ID NOS: 3381

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2223

; LENGTH: 116

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-094-749-2223

Query Match 0.8%; Score 7; DB 12; Length 116;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 357 AAAGSA 363
| | | | |
Db 15 AAAGSA 21

RESULT 40

US-09-739-907-182
; Sequence 182, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/739,907
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 05/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 182
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-739-907-182

Query Match 0.8%; Score 7; DB 9; Length 118;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 357 AAAGSA 363
| | | | |
Db 82 AAAGSA 88

Search completed: February 17, 2004, 10:26:52
Job time : 51 secs

THIS PAGE BLANK (USPTO)

Db 408 GFELANQVGNITKAVSYIIAQRVAAGLSTGVPVALLIASTVSLAISPLAFAGIADKFN 467
 Qy 305 HANALDEFKQFRKFGYDGDHLLAEYQGVGTIEASLTITSTALGAVSAGVAAAAGSVA 364
 Db 468 HAKLESYAERFKGLGYDGNLLAEYQGVGTIDASVTAINALAAAGVSAAGSVI 527
 Qy 365 GTPIALVAGVTGLISGLEASQAMFESVANRLOGLKEWEKONGQNYFDKGYDSRYA 424
 Db 528 ASPIALVSGITGVISTILQYSQAMFEHVANKHNKIVEWKNHKNKYNFENGIDARYL 587
 Qy 425 AYLANNLKFLSELNKELEAEERVAITTOQRWNNIGELAGITKLGERIKSKAYADAFEDG 484
 Db 588 ANLQDNMKFLNLNKLQAEERVAITTOQRWNNIGELAGISRLGKESKAYADAFEEG 647
 Qy 485 KYVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTSRRLTNGKYSYINKLK 544
 Db 648 KHAKADKLVLQDSANGIIDVSNNGKKTQHILPRTPLTPGTEHRRVQGVGYEYITKLN 707
 Qy 545 FGRVKNQVTDGASSKLDFSKVIQV-----AETGDEGLIIVNAGNDDIFVQ 597
 Db 708 INRVDSWKITDGAASSTFDLTVVQRIIGELDNAGNVTTKTKIIAKLGECDNDFVGS 767
 Qy 598 GKNNIDGGHDRVFSYKDGKGNITVDGTSATEAGSYTVNRKVARGDIVHEVVKQETK 657
 Db 768 GTTEIDGGEGYDRVHSR-GNYGALTIDATKETEQQSYTVNRVFTGKALHEVTSHTAL 826
 Qy 658 VGRVKNQVTDGASSKLDFSKVIQV-----AETGDEGLIIVNAGNDDIFVQ 597
 Db 827 VGNREEKIEYR-HSNQNHAGYVTKOTLKAVEBIICTSHNDIFKGSKEADNANGGDDVT 885
 Qy 718 LDGAGDDBLFGKGNDRSLGDDLLDGGSDVVLNGAGNDVYIFKGDNDTLYDG 777
 Db 886 IDGNDGNLFLFGKGNDRSLGDDLLDGGSDVVLNGAGNDVYIFKGDNDTLYDG 777
 Qy 778 TGNDKLAFADANISIMIERTEGIIIVKNDHSGSINIPRW-----ITSNLQYOSNKT 833
 Db 946 DGNDKLSFSDSNLKDITFEKVXNLVI-TNSKKEVTIQWTFREADFAKEVFNKATK-D 1003
 Qy 834 HKIEQLIGKGYTSDDQDKLQDKOGTVITSQELKKLADENKSKLSASDIASSLNK 893
 Db 1004 EKIEEIIQNGRITSKQVDDLI--AKNGKITQDELSKVVDNVELLKH-S-KNVVNSLDK 1060
 Qy 894 LVGSMALFTANSVSNALQIPITQPTOGI 922
 Db 1061 LISSVSAFTSSNDSRNLVAPTSMLDQSL 1089

RESULT 4

AAR22103
 AC AAR22103 standard; Protein; 1098 AA.
 AC AAR22103;
 DT 06-JUL-1992 (first entry)
 DE Bovine IL-2 - LKT fusion protein.
 KW Interleukin 2; leuko-toxin; vaccine; pneumonia; respiratory diseases.
 OS Pasteurella haemolytica.
 OS Bos taurus.
 XX WO9203558-A.
 XX 05-MAR-1992.
 PD 22-AUG-1991; 91WO-CA00299.
 XX 22-AUG-1990; 90US-0571301.
 XX (POTT)/ POTTER A.

PI Potter A, Campos M, Hughes HPA;
 XX WPI: 1992-096901/12.
 DR N-PSDB; AAO22771.
 XX Interleukin 2-leuko-toxin gene fusion - encodes fusion protein useful as vaccine for animal pneumonia
 PT Claim 20; Fig 3; 68pp; English.
 PS
 XX The IL-2-LKT protein was encoded by a chimeric gene contg. the bovine IL-2 gene fused to the DNA encoding at least one epitope of leukotoxin from *P. haemolytica*. IL-2-LKT was gel purified and ligated into the expression vector pGH433 lacI. The resulting clone pAA356 (ATCC 68396) contd. the desired gene fusion under the control of the *E. coli* lac promoter. The protein produced by the gene fusion is useful in a vaccine compn. with a pharmaceutically acceptable vehicle, e.g. a carrier homologous to a rotavirus VP6 inner capsid protein. The vaccine can be used for preventing or ameliorating respiratory diseases in animals e.g. shipping fever or pneumonia.
 CC See also AAR24124, 5.
 CC
 XX Sequence 1098 AA;
 SQ

Query Match 50.2%; Score 2332; DB 13; Length 1098;
 Best Local Similarity 50.1%; Pred. No. 9.9e-137;
 Matches 465; Conservative 174; Mismatches 264; Indels 26; Gaps 13;
 Qy 8 KSNIOAGLNSTKSLGNLYLAIPKD--YDPKQGTLDNDFIKAADELGIARLAEPNHTET 65
 Db 173 QSLTOAG--SSLTKCAKIIILYIPQYDYDEQNGLDLVKAAEELGIEVQREERNIAT 231
 Qy 66 AKKSVDTWQFLSTQGTIAISATKLEKFLQKSTNKLAKGLDSVENIDRKLKASNVLS 125
 Db 232 AOTSGLGTQTAIGTERGIVLSAPQIDKLQK---TRAGQALGSAESIVQNAKAKTVLS 288
 Qy 126 TLSSFLGTALAGIELDLSIKKGAAPDALAKASITDLNIEIIGNLSQSTQTIETAFSSQAK 185
 Db 289 GIQSLGSVLGMDLDEL--QNNNSQHALAKAGLELNTSLIENIANSVKTLDEPGEQISQ 347
 Qy 186 LGSTISQAKGFSNIGNKQNL-NFSKTNLGLIEITGLLSGISAGFALADNASTGKKVAA 244
 Db 348 FGSKLQNIKGLTGLDKLKNIGGLDKAGIGLDVISGLSGATAALVLADKNASTAKVGA 407
 Qy 245 GFELSNQVIGNVTKATISSYLAQRAAGLSTGCAVAALITSSIMLAISPLAFNADKFN 304
 Db 408 GFELANQVGNITKAVSYIIAQRVAAGLSTGVPVALLIASTVSLAISPLAFAGIADKFN 467
 Qy 305 HANALDEFKQFRKFGYDGDHLLAEYQGVGTIEASLTITSTALGAVSAGVAAAAGSVA 364
 Db 468 HAKLESYAERFKGLGYDGNLLAEYQGVGTIDASVTAINALAAAGVSAAGSVI 527
 Qy 365 GTPIALVAGVTGLISGLEASQAMFESVANRLOGLKEWEKONGQNYFDKGYDSRYA 424
 Db 528 ASPIALVSGITGVISTILQYSQAMFEHVANKHNKIVEWKNHKNKYNFENGIDARYL 587
 Qy 425 AYLANNLKFLSELNKELEAEERVAITTOQRWNNIGELAGITKLGERIKSKAYADAFEDG 484
 Db 588 ANLQDNMKFLNLNKLQAEERVAITTOQRWNNIGELAGISRLGKESKAYADAFEEG 647
 Qy 485 KYVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTSRRLTNGKYSYINKLK 544
 Db 648 KHAKADKLVLQDSANGIIDVSNNGKKTQHILPRTPLTPGTEHRRVQGVGYEYITKLN 707
 Qy 545 FGRVKNQVTDGASSKLDFSKVIQV-----AETGDEGLIIVNAGNDDIFVQ 597
 Db 708 INRVDSWKITDGAASSTFDLTVVQRIIGELDNAGNVTTKTKIIAKLGECDNDFVGS 767
 Qy 598 GKNNIDGGHDRVFSYKDGKGNITVDGTSATEAGSYTVNRKVARGDIVHEVVKQETK 657
 Db 768 GTTEIDGGEGYDRVHSR-GNYGALTIDATKETEQQSYTVNRVFTGKALHEVTSHTAL 826

THIS PAGE BLANK (USPTO)

Db 173 QSLTQAG-SSLTKGAKKIYIPONYQYDTEQNGLODLVKAABEIGIEVOREERNIAT 231
 QY 66 AKKSVDTVNOFLSLTGTGIAISATKLEKFLQKHSSTNKLAAGLSDVENIDRKLKASNVLS 125
 Db 232 AQTSLGTIQTALGTERGIVLSAPOIDKLLQK---TKAQALGSAESIIVONANKAKTVLS 288
 QY 126 TLSSFLGTALAGIELDSLKKGAADAPALAKASIDLINEIGNLSQSTOTIEAFSSQLAK 185
 Db 289 GIQSILGSLVAGMDLDEAL-QNNSNOHALAKAGLELTNSLIENIANSVKTLDFEQEIQS 347
 QY 186 LGSTISQAKGFSNIGKQLNL-NFSKTNLGLLEITGLLSGISAGPALADKNASTGKVA 244
 Db 348 FGSKLQNIKGLTGLDKLKNIGGLDAGLDVIGLLSGATAAVLADKNASTAKKVA 407
 QY 245 GFELSNQVGNVTKATSSYVLAORVAAGLSTTGAVAAITSSIMLAISPLAFNNAADKFN 304
 Db 408 GFELANQVGNVTKAVSSYVLAORVAAGLSTTGAVAAITSSIMLAISPLAFIADKEN 467
 QY 305 HANALDEFAKFRKFGYDGHLLAEYORGVTTEASLTITSTALGAVSAGVSAAGVSAV 364
 Db 468 HAKSLESYAEFRKGLGYDGNLLAEYORGVTTEASLTITSTALGAVSAGVSAAGVSAV 527
 QY 365 GTPALAVAGVTGLIGLEASQAMPESVANRLQKLEWEKONGSONYFDKGYDSRYA 424
 Db 528 ASPIALVSGTIVSTILQYSKOAMPEHVAHNKIVWEKNHKGKNFENGIDARYL 597
 QY 425 AYLANNLKFLSELNKELEAERVAITAOQRDNNIGELAGITKLGERIKSGKAYADAFEDG 484
 Db 588 ANLQDNKFFLLNKELEAERVAITAOQRDNNIGELAGITKLGERIKSGKAYADAFEDG 647
 QY 485 KQVEAGSNITLDKTIIDISNGKKTQALHPTSPILLTAGTSERELTNGKYSYINKLK 544
 Db 648 KHAKADKLVLDSANGIIDVSNKGAKTQHILFRTPLTPTGTEHRRVQKGYEYITKLN 707
 QY 545 FGRVKNQVQDGGASSKLDKSKVIQV-----AETEGTDEIGLIVNAKAGNDIDFVGG 597
 Db 708 INRVDSWQDGGASSKLDKSKVIQV-----AETEGTDEIGLIVNAKAGNDIDFVGG 767
 QY 598 GKNWIDGGDHRVYFYSKDGFGNITVDGTSATEAGSYVNRKVRGDNVHEVVKRQETK 657
 Db 768 GTTEIDGEGYDRVHYSR-GNYGALTIDATKETEFGSSYTVNRVETGKALHEVTSHTAL 826
 QY 658 VGRKTETIQVRDELKRVGYQSTDMKSVVEVIGSQFNDVPKSGKFNDIFHSGEGDDL 717
 Db 827 VGNREKIEYR-HSNQNHAGYTYKTOLKAVEEIIIGTSHNDIFKSGKFNDAFNGDGVDT 885
 QY 718 LQGAGDRLFGKGNLSDGDDLLDGGSGDDVLLGNGNDVYIFRKGDDNDTLVDG 777
 Db 886 IDGNDGNDRLFGKGNLSDGDDLLDGGSGDDVLLGNGNDVYIFRKGDDNDTLVDG 945
 QY 778 TGNDKLAFAADANISDIMIERTKEGIIVRNDHSGSINIPRWY---ITSNLQYQSNKTD 833
 Db 946 DENDKLSFSDSLKDLTFEKVKNLVI-TSKKEKVITQWFEADPAKEVPVNAKATK-D 1003
 QY 834 HKIEQLKQGSYTSQIDKLODKGDTTTSQELKULADENKSKLSASDIASLXK 893
 Db 1004 EKIEBIIICONGERITTSQVDDLEA-AKNGKTIODELSKVVDNYELLKHS-KNVTNSLDK 1060
 QY 894 LVGSMALFGTANSVSSNALOPTOPTQGI 922
 Db 1061 LISSVAFSTNSDRNVVAPTSMLDQSL 1089

RESULT 3
 AAB21073
 ID AAB21073 standard; Protein, 1098 AA.
 XX AC AAB21073;
 XX DT 19-DEC-2000 (first entry)
 XX DE Bovine IL-2/Pasteurella haemolytica leukotoxin fusion protein.

XX KW Bovine IL-2; interleukin-2; leukotoxin; LKT; respiratory disease; pneumonia; shipping fever; cattle; livestock; anti-Pasteurella vaccine; immunogen.
 XX OS Chimeric - Bos taurus.
 XX OS Chimeric - Pasteurella haemolytica.
 XX OS US6096920-A.
 XX PR 20-OCT-1997; 97US-0954418.
 XX PR 20-DEC-1993; 93US-0170126.
 XX PR 22-JUL-1996; 96US-0681479.
 XX PR 22-AUG-1990; 90US-0571301.
 XX PR 16-OCT-1991; 91US-0777715.
 XX PA (UYSA-) UNIV SASKATCHEWAN.
 XX PA (CIBA) CIBA GEIGY CANADA LTD.
 XX PR Campbell M.; Hughes HPA, Potter A;
 XX DR WPI: 2000-531543/48.
 XX DR N-PSDB; AAA72483.
 XX PS Vaccine for stimulating immunity against pneumonia comprises chimeric protein comprising gamma-interferon and leukotoxin derived from Pasteurella haemolytica
 XX Example 1; Column 31-38; 56pp; English.
 CC The invention relates to a novel vaccine composition comprising an immunogenic chimeric protein that comprises gamma-interferon (gamma-IFN) or an active fragment thereof, linked to an epitope of a Pasteurella haemolytica leukotoxin (LKT). Pasteurella species, especially Pasteurella haemolytica, are responsible for respiratory diseases in a range of agricultural animals, most particularly cattle, but also sheep, pigs, horses and fowl. Shipping fever is the most economically important respiratory disease associated with Pasteurella species, affecting 15-30% of exposed cattle and resulting in a 2-5% mortality rate in the exposed population. The vaccine composition of the invention is useful for preventing or ameliorating respiratory diseases such as pneumonia, particularly shipping fever pneumonia, in livestock. The present sequence represents a fusion protein comprising bovine interleukin-2 (IL-2) and Pasteurella haemolytica leukotoxin, which may also be used as an anti-Pasteurella vaccine.

SQ Sequence 1098 AA;

Query Match 50.2%; Score 2334; DB 21; Length 1098;

Best Local Similarity 50.1%; Pred. No. 7.4e-137;

Matches 465; Conservative 175; Mismatches 263; Indels 26; Gaps 13;

QY 8 KSNICAGLNTSKGLNLYLAIPKD--YDPQGGTINDFIKADELGIARLAEENHTT 65

Db 173 QSLTQAG-SSLTKGAKKIYIPONYQYDTEQNGLODLVKAABEIGIEVOREERNIAT 231

QY 66 AKKSVDTVNOFLSLTGTGIAISATKLEKFLQKHSSTNKLAAGLSDVENIDRKLKASNVLS 125

Db 232 AQTSLGTIQTALGTERGIVLSAPOIDKLLQK---TKAQALGSAESIIVONANKAKTVLS 288

QY 126 TLSSFLGTALAGIELDSLKKGAADAPALAKASIDLINEIGNLSQSTOTIEAFSSQLAK 185

Db 289 GIQSILGSLVAGMDLDEAL-QNNSNOHALAKAGLELTNSLIENIANSVKTLDFEQEIQS 347

QY 186 LGSTISQAKGFSNIGKQLNL-NFSKTNLGLLEITGLLSGISAGPALADKNASTGKVA 244

Db 348 FGSKLQNIKGLTGLDKLKNIGGLDAGLDVIGLLSGATAAVLADKNASTAKKVA 407

QY 245 GFELSNQVGNVTKATSSYVLAORVAAGLSTTGAVAAITSSIMLAISPLAFNNAADKFN 304

THIS PAGE BLANK (USPTO)

Baskar, P.
10/069799
Text Search

10/069799

FILE 'HCAPLUS, MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH,
JICST-EPLUS, JAPIO' ENTERED AT 09:57:43 ON 18 FEB 2004

L1 24 S "FARN J"?/AU
L2 390 S "STRUGNELL R"?/AU
L3 186 S "TENNENT J"?/AU
L4 7 S L1 AND L2 AND L3
L5 8 S L1 AND (L2 OR L3)
L6 7 S L2 AND L3
L7 49 S (L1 OR L2 OR L3) AND BOVIS
L8 50 S L4 OR L5 OR L6 OR L7
L9 13 DUP REM L8 (37 DUPLICATES REMOVED)

- Author(s)

L9 ANSWER 1 OF 13 HCAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 1
ACCESSION NUMBER: 2001:168028 HCAPLUS
DOCUMENT NUMBER: 134:221433
TITLE: Vaccine antigens of Moraxella
INVENTOR(S): Farn, Jacinta; Strugnell,
Richard; Tennent, Jan
PATENT ASSIGNEE(S): Commonwealth Scientific and Industrial Research
Organisation, Australia; The University of
Melbourne
SOURCE: PCT Int. Appl., 60 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

| PATENT NO. | KIND | DATE | APPLICATION NO. | DATE |
|------------------------|--|----------|-----------------|------------|
| WO 2001016172 | A1 | 20010308 | WO 2000-AU1048 | 20000831 |
| W: | AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM | | | |
| RW: | GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG | | | |
| EP 1210364 | A1 | 20020605 | EP 2000-955974 | 20000831 |
| R: | AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL | | | |
| BR 2000013574 | A | 20020611 | BR 2000-13574 | 20000831 |
| PRIORITY APPLN. INFO.: | | | AU 1999-2571 | A 19990831 |
| | | | WO 2000-AU1048 | W 20000831 |

AB The present invention relates to antigens of Moraxella, in particular, Moraxella bovis, nucleic acid sequences encoding these antigens and formulations for use in raising an immune response against Moraxella.

REFERENCE COUNT: 5 THERE ARE 5 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 2 OF 13 HCAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 2
ACCESSION NUMBER: 2001:811555 HCAPLUS
DOCUMENT NUMBER: 136:66038

Searcher : Shears 571-272-2528

THIS PAGE BLANK (USPTO)

10/069799

TITLE: Molecular characterization of a secreted enzyme with phospholipase B activity from *Moraxella bovis*
AUTHOR(S): Farn, Jacinta L.; Strugnell, Richard A.; Hoyne, Peter A.; Michalski, Wojtek P.; Tennent, Jan M.
CORPORATE SOURCE: CSIRO Livestock Industries, Geelong, 3220, Australia
SOURCE: Journal of Bacteriology (2001), 183(22), 6717-6720
CODEN: JOBAAY; ISSN: 0021-9193
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB A candidate for a vaccine against infectious bovine keratoconjunctivitis (IBK) has been cloned and characterized from *Moraxella bovis*. The plb gene encodes a protein of 616 amino acids (mol. mass of .apprx.65.8 kDa) that expresses phospholipase B activity. Amino acid sequence anal. revealed that PLB is a new member of the GDSL (Gly-Asp-Ser-Leu) family of lipolytic enzymes.

REFERENCE COUNT: 32 THERE ARE 32 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 3 OF 13 HCAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 3

ACCESSION NUMBER: 2000:359019 HCAPLUS
DOCUMENT NUMBER: 133:100763
TITLE: Characterization of hemolysin of *Moraxella bovis* using a hemolysis-neutralizing monoclonal antibody
AUTHOR(S): Billson, F. Mark; Harbour, Colin; Michalski, Wojtek P.; Tennent, Jan M.; Egerton, John R.; Hodgson, Jennifer L.
CORPORATE SOURCE: Department of Veterinary Clinical Sciences, University of Sydney, Camden, 2570, Australia
SOURCE: Infection and Immunity (2000), 68(6), 3469-3474
CODEN: INFIBR; ISSN: 0019-9567
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB A concentrated bacterial culture supernatant from the hemolytic *Moraxella bovis* strain UQV 148NF was used to immunize mice and generate monoclonal antibodies (MAbs). One, MAb G3/D7, neutralized the hemolytic activity of *M. bovis* and recognized a 94-kDa protein by Western blot anal. in hemolytic *M. bovis* strains representing each of the different fimbrial serogroups. Exposure of corneal epithelial cells to *M. bovis* concentrated culture supernatants demonstrated a role for an exotoxin in the pathogenesis of infectious bovine keratoconjunctivitis, while neutralization of hemolytic and cytotoxic activities by MAb G3/D7 implies that these activities are related or have common epitopes. The action of *M. bovis* hemolysin was further characterized in sheep erythrocyte preps. with a binding step and Ca²⁺ required for lysis to proceed, similar to the RTX family of bacterial exotoxins. Neutralization of lytic activity in vitro is evidence for the presence of *M. bovis* antigens, which may be capable of protecting cattle from the development of infectious

THIS PAGE BLANK (USPTO,

10/069799

bovine keratoconjunctivitis.

REFERENCE COUNT: 39 THERE ARE 39 CITED REFERENCES AVAILABLE
FOR THIS RECORD. ALL CITATIONS AVAILABLE
IN THE RE FORMAT

L9 ANSWER 4 OF 13 HCAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 4
ACCESSION NUMBER: 2000:507582 HCAPLUS
DOCUMENT NUMBER: 133:249416
TITLE: Identification of type 4 fimbriae in
Actinobacillus pleuropneumoniae
AUTHOR(S): Zhang, Y.; Tennent, J. M.; Ingham, A.;
Beddome, G.; Prideaux, C.; Michalski, W. P.
CORPORATE SOURCE: Private Bag 24, Australian Animal Health
Laboratory, CSIRO Animal Health, Geelong, 3220,
Australia
SOURCE: FEMS Microbiology Letters (2000), 189(1), 15-18
CODEN: FMLED7; ISSN: 0378-1097
PUBLISHER: Elsevier Science B.V.
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Type 4 fimbriae have been identified on the cell surface of
Actinobacillus pleuropneumoniae by electron microscopy and
N-terminal sequencing anal. A. pleuropneumoniae type 4 fimbrial
subunit protein, purified from cell cultures and from outer membrane
prepns., reacted with polyclonal antibody raised against type 4
fimbriae of Moraxella bovis on Western blots. N-terminal
sequence anal. of the purified 17 kDa type 4 fimbrial subunit
protein, named ApfA, revealed the first 12 amino acids to be
identical to those of other type 4 fimbrial subunit proteins.

REFERENCE COUNT: 23 THERE ARE 23 CITED REFERENCES AVAILABLE
FOR THIS RECORD. ALL CITATIONS AVAILABLE
IN THE RE FORMAT

L9 ANSWER 5 OF 13 HCAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 5
ACCESSION NUMBER: 1997:19161 HCAPLUS
DOCUMENT NUMBER: 126:57234
TITLE: Identification, purification, and
characterization of the type 4 fimbriae of
Pasteurella multocida
AUTHOR(S): Ruffolo, Carmel G.; Tennent, Jan M.;
Michalski, Wojtek P.; Adler, Ben
CORPORATE SOURCE: Dep. Microbiology, Monash Univ., Clayton, 3168,
Australia
SOURCE: Infection and Immunity (1997), 65(1), 339-343
CODEN: INFIBR; ISSN: 0019-9567
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The presence of fimbriae on Pasteurella multocida has been reported,
but there have been no prior studies aimed at conclusively
characterizing these structures. We now report on the
identification and characterization of type 4 fimbriae on serogroup
A, B, and D strains of P. multocida. Under microaerophilic
conditions P. multocida showed an increased expression of the
fimbriae, which were observed to form bundles. Fimbriae purified by
high-performance reverse-phase liquid chromatog. constituted a single
18-kDa subunit, the first 21 amino acids of which shared very high
similarity with the N-terminal amino acids sequence of other type 4

THIS PAGE BLANK (USPTO)

10/069799

fimbrial subunits. Antiserum against the *P. multocida* 18-kDa protein immunostained the type 4 fimbrial subunit of *Moraxella bovis* and *Dichelobacter nodosus*. Based on these observations we conclude that *P. multocida* possesses type 4 fimbriae and have designated the *P. multocida* fimbrial subunit Ptf/A.

L9 ANSWER 6 OF 13 BIOSIS COPYRIGHT 2004 BIOLOGICAL ABSTRACTS INC. on STN

ACCESSION NUMBER: 1996:399092 BIOSIS
DOCUMENT NUMBER: PREV199699121448
TITLE: Construction of a defined aromatic mutant of the gram-positive bacterium *C. pseudotuberculosis*.
AUTHOR(S): Simmons, Cameron [Reprint author]; Farn, Jacinta [Reprint author]; Strugnell, Richard [Reprint author]; Hodgson, Adrian
CORPORATE SOURCE: Dep. Microbiol., Univ. Melbourne, Parkville, VIC 3052, Australia
SOURCE: Brown, F. [Editor]; Norrby, E. [Editor]; Burton, D. [Editor]; Mekalanos, J. [Editor]. Vaccines (Cold Spring Harbor), (1996) pp. 111-116. Vaccines (Cold Spring Harbor); Molecular approaches to the control of infectious diseases.
Publisher: Cold Spring Harbor Laboratory Press, 10 Skyline Drive, Plainview, New York 11803, USA.
Series: Vaccines (Cold Spring Harbor).
Meeting Info.: Thirteenth Meeting. Cold Spring Harbor, New York, USA. September 13-17, 1995.
ISSN: 0899-4056. ISBN: 0-87969-479-3.
DOCUMENT TYPE: Book
Conference; (Meeting)
Book; (Book Chapter)
Conference; (Meeting Paper)
LANGUAGE: English
ENTRY DATE: Entered STN: 3 Sep 1996
Last Updated on STN: 3 Sep 1996

L9 ANSWER 7 OF 13 MEDLINE on STN DUPLICATE 6

ACCESSION NUMBER: 95336598 MEDLINE
DOCUMENT NUMBER: PubMed ID: 7542003
TITLE: Antibody reactivity to mycobacterial 65 kDa heat shock protein: relevance to autoimmunity.
AUTHOR: Karopoulos C; Rowley M J; Handley C J; Strugnell R A
CORPORATE SOURCE: Department of Biochemistry, Monash University, Clayton, Victoria, Australia.
SOURCE: Journal of autoimmunity, (1995 Apr) 8 (2) 235-48.
Journal code: 8812164. ISSN: 0896-8411.
PUB. COUNTRY: ENGLAND: United Kingdom
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199508
ENTRY DATE: Entered STN: 19950905
Last Updated on STN: 19960129
Entered Medline: 19950818
AB Reactivity to the mycobacterial 65 kDa heat shock protein (HSP 65) has been implicated in the pathogenesis of adjuvant arthritis in the rat, and may be involved in the pathogenesis of rheumatoid arthritis

Searcher : Shears 571-272-2528

THIS PAGE BLANK (USPTO)

or other autoimmune diseases in humans. Accordingly this study sought quantitative or qualitative differences in the antibody reactivity to HSP 65 between normal controls, patients with the multisystem autoimmune diseases, rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) and patients with the mycobacterial infections, tuberculosis (TB) and leprosy. Levels of antibodies to recombinant HSP 65 in serum were measured by ELISA in normal subjects and in patients with RA, SLE, TB or leprosy. Antibody reactivity was examined by Western blotting using polypeptide fragments of HSP 65 derived by recombinant DNA techniques, or by digestion with trypsin or cyanogen bromide (CNBr). Reactivity to a synthetic peptide, the adjuvant arthritis T-cell epitope of HSP 65 (180-188), was tested by ELISA. High levels of antibodies to full length recombinant HSP 65 from *Mycobacterium bovis* were present in all the groups tested. By Western blot analysis, most reactivity with intact HSP 65 was retained in a 32 kDa tryptic fragment, judged by sequencing and size estimations to represent amino acid residues 118- approximately 388. This sequence included a major T-cell epitope for adjuvant arthritis (180-188), but these nine amino acids were not essential for B-cell reactivity since most sera also reacted with residues 188-540 which lack the T-cell epitope. Moreover, the 180-188 synthetic peptide was unreactive by ELISA, and did not inhibit reactivity with the intact recombinant HSP 65. In conclusion, most individuals had antibodies to mycobacterial HSP 65, presumably resulting from previous bacterial infections. The magnitude of the response was unrelated to the occurrence of systemic autoimmune disease, and the pattern of antibody reactivity with recombinant and proteolytic fragments of HSP 65 suggests that the major B-cell epitope is conformational and consists of discontinuous regions of the molecule.

L9 ANSWER 8 OF 13 MEDLINE on STN DUPLICATE 7
 ACCESSION NUMBER: 96037961 MEDLINE
 DOCUMENT NUMBER: 96037961 PubMed ID: 7571364
 TITLE: The protective efficacy of cloned *Moraxella bovis* pili in monovalent and multivalent vaccine formulations against experimentally induced infectious bovine keratoconjunctivitis (IBK).
 AUTHOR: Lepper A W; Atwell J L; Lehrbach P R; Schwartzkoff C L; Egerton J R; Tennent J M
 CORPORATE SOURCE: CSIRO Division of Animal Health, Animal Health Research Laboratory, Parkville, Vic., Australia.
 SOURCE: VETERINARY MICROBIOLOGY, (1995 Jul) 45 (2-3) 129-38. Journal code: 7705469. ISSN: 0378-1135.
 PUB. COUNTRY: Netherlands
 DOCUMENT TYPE: (CLINICAL TRIAL)
 (CONTROLLED CLINICAL TRIAL)
 Journal; Article; (JOURNAL ARTICLE)
 LANGUAGE: English
 FILE SEGMENT: Priority Journals
 ENTRY MONTH: 199511
 ENTRY DATE: Entered STN: 19951227
 Last Updated on STN: 19951227
 Entered Medline: 19951108
 AB Calves were vaccinated with cloned *Moraxella bovis* pili of serogroup C (experiment 1) or B (experiment 2) either as a monovalent formulation or as part of a multivalent preparation with

THIS PAGE BLANK (USPTO)

pili of six other serogroups. Within 4 weeks of the second vaccine dose vaccinated calves and non-vaccinated controls were challenged via the ocular route with either virulent *M. bovis* strain Dal2d (serogroup C) or *M. bovis* strain 3W07 (serogroup B) in experiments 1 and 2, respectively. Calves vaccinated with multivalent vaccines had significantly lower antibody titres than those vaccinated with monovalent preparations. Nevertheless, the levels of protection against infectious bovine keratoconjunctivitis (IBK) achieved with multivalent vaccines were 72% and 83% for the groups challenged with *M. bovis* strains of serogroups B and C, respectively. The serogroup C monovalent vaccine gave 100% protection against experimentally induced IBK and *M. bovis* isolates cultured from the eyes 6 days post-challenge were identified as belonging solely to serogroup C. Unexpectedly, only 25% protection was achieved against homologous strain challenge of calves that received the monovalent serogroup B vaccine. Furthermore, the majority of *M. bovis* isolates recovered from calves in this group belonged to serogroup C, as did half of those isolates cultured from the multivalent vaccinates. The remaining bacterial isolates from the latter group, together with all isolates from the non-vaccinated controls, belonged to serogroup B. Results are consistent with the hypothesis that derivatives of the serogroup B challenge inoculum had expressed serogroup C pilus antigen within 6 days of the challenge, possibly as a result of pilus gene inversion occurring in response to the presence of specific antibody in eye tissues and tears.

L9 ANSWER 9 OF 13 HCAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 8
 ACCESSION NUMBER: 1994:694086 HCAPLUS
 DOCUMENT NUMBER: 121:294086
 TITLE: Characterization of pilin genes from seven serologically defined prototype strains of *Moraxella bovis*
 AUTHOR(S): Atwell, John L.; Tennent, Jan M.; Lepper, Anthony W. D.; Elleman, Tom C.
 CORPORATE SOURCE: Commonwealth Scientific and Industrial Research Organisation, Victoria, 3052, Australia
 SOURCE: Journal of Bacteriology (1994), 176(16), 4875-82
 CODEN: JOBAAY; ISSN: 0021-9193
 DOCUMENT TYPE: Journal
 LANGUAGE: English

AB Numerous field isolates of *Moraxella bovis* have previously been classified by serol. techniques into seven serogroups, each defined by homologous cross-reaction with antisera prepared against purified pili of a single prototype strain. The gene encoding pilin from each of the prototype strains has been characterized by nucleotide sequence determination. The coding sequences show extensive homol. (70 to 80%) while the strains were also characterized. The presence of an addnl., partial pilin gene in each prototype strain was confirmed by Southern blot anal., and the partial pilin genes from two strains of one serogroup were characterized by sequence determination. Features of the pilin gene sequences are considered in relation to pilin gene inversion and the serol. variants which may arise from inversion events.

L9 ANSWER 10 OF 13 HCAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 9
 ACCESSION NUMBER: 1995:236681 HCAPLUS
 TITLE: A haemolytic cell-free preparation of *Moraxella*

THIS PAGE BLANK (USPTO)

10/069799

bovis confers protection against
Infectious Bovine Keratoconjunctivitis

AUTHOR(S): Billson, F. Mark; Hodgson, Jennifer L.; Egerton,
John R.; Lepper, Anthony W. D.; Michalski,
Wojtek P.; Schwartzkoff, C. Leigh; Lehrbach,
Philip R.; Tennent, Jan M.

CORPORATE SOURCE: University of Sydney, Department of Animal
Health, Private Bag 3, Camden, NSW, 2570,
Australia

SOURCE: FEMS Microbiology Letters (1994), 124(1), 69-74
CODEN: FMLED7; ISSN: 0378-1097

PUBLISHER: Elsevier

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Protection conferred by a cell-free preparation from a haemolytic
Moraxella **bovis** isolate, UQV 148NF, was compared to an
equivalent fraction from a non-haemolytic M. **bovis** isolate,
Gordon 26L3, and to a recombinant DNA-derived pili vaccine. Three
groups of ten calves were vaccinated twice with one of the three
preps. and, together with ten non-vaccinated calves, challenged
with virulent M. **bovis** isolate Dal 2d. Compared to the
control group, significant protection was observed in the group
receiving the pili vaccine and the group receiving the preparation from
haemolytic isolate, UQV 148NF.

L9 ANSWER 11 OF 13 MEDLINE on STN DUPLICATE 10

ACCESSION NUMBER: 94055030 MEDLINE

DOCUMENT NUMBER: 94055030 PubMed ID: 7901935

TITLE: A Moraxella **bovis** pili vaccine produced by
recombinant DNA technology for the prevention of
infectious bovine keratoconjunctivitis.

AUTHOR: Lepper A W; Elleman T C; Hoyne P A; Lehrbach P R;
Atwell J L; Schwartzkoff C L; Egerton J R;
Tennent J M

CORPORATE SOURCE: CSIRO Division of Animal Health, Animal Health
Research Laboratory, Parkville, Vic., Australia.

SOURCE: VETERINARY MICROBIOLOGY, (1993 Jul) 36 (1-2) 175-83.
Journal code: 7705469. ISSN: 0378-1135.

PUB. COUNTRY: Netherlands

DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)

LANGUAGE: English

FILE SEGMENT: Priority Journals

ENTRY MONTH: 199312

ENTRY DATE: Entered STN: 19940117
Last Updated on STN: 19970203
Entered Medline: 19931203

AB Pili (fimbriae) were prepared from Moraxella **bovis** strain
Dalton 2d (Dal2d) and from a derivative of Pseudomonas aeruginosa
K/2Pfs that contained a plasmid-borne Dal2d pilin gene and produced
pili having serogroup-specific identity to Dal2d. Nine calves were
vaccinated with two doses each of 30 micrograms authentic M.
bovis Dal2d pili in oil adjuvant and 10 calves were
vaccinated with a similar dose of P. aeruginosa-derived Dal2d pili
in the same formulation. All 19 calves and 10 non-vaccinated
controls were challenged by instillation of 1×10^9 virulent M.
bovis Dal2d cells into both conjunctival sacs 19 days after
the second vaccine dose. The serological response to vaccination
and the degree of protection against experimentally induced

THIS PAGE BLANK (USPTO)

infectious bovine keratoconjunctivitis (IBK) were assessed. None of the nine calves vaccinated with authentic *M. bovis* Dal2d pili developed IBK while two of those vaccinated with *P. aeruginosa*-derived Dal2d pili developed lesions which accounted for a mean group lesion score of 0.3. In contrast, 9 of the 10 non-vaccinated calves developed IBK lesions, the majority of which were progressive, required early treatment and accounted for a mean group lesion score of 1.5. These results demonstrate the potential of a relatively low dose of pili produced by recombinant DNA technology for development of an effective vaccine against IBK.

L9 ANSWER 12 OF 13 BIOSIS COPYRIGHT 2004 BIOLOGICAL ABSTRACTS INC. on STN

ACCESSION NUMBER: 1993:200425 BIOSIS
DOCUMENT NUMBER: PREV199344096675
TITLE: Purification of proteases by immobilised bacitracin A affinity chromatography.
AUTHOR(S): Michalski, Wojtek P.; Crooks, Jennipher K.; Prowse, Stephen J.; Tennent, Jan M.; Lepper, Anthony W. D.
CORPORATE SOURCE: CSIRO Div. Anim. Health, Anim. Health Res. Lab., Private Bag No. 1, Parkville, Victoria 3052, Australia
SOURCE: Journal of Cellular Biochemistry Supplement, (1993) Vol. 0, No. 17 PART A, pp. 50.
Meeting Info.: Keystone Symposium on Protein Purification and Biochemical Engineering. Santa Fe, New Mexico, USA. January 15-21, 1993.
ISSN: 0733-1959.
DOCUMENT TYPE: Conference; (Meeting)
LANGUAGE: English
ENTRY DATE: Entered STN: 16 Apr 1993
Last Updated on STN: 9 Jun 1993

L9 ANSWER 13 OF 13 MEDLINE on STN DUPLICATE 11

ACCESSION NUMBER: 93069932 MEDLINE
DOCUMENT NUMBER: 93069932 PubMed ID: 1359693
TITLE: The protective efficacy of pili from different strains of *Moraxella bovis* within the same serogroup against infectious bovine keratoconjunctivitis.
AUTHOR: Lepper A W; Moore L J; Atwell J L; Tennent J M
CORPORATE SOURCE: CSIRO Division of Animal Health, Animal Health Research Laboratory, Parkville, Vic., Australia.
SOURCE: VETERINARY MICROBIOLOGY, (1992 Sep) 32 (2) 177-87.
Journal code: 7705469. ISSN: 0378-1135.
PUB. COUNTRY: Netherlands
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199212
ENTRY DATE: Entered STN: 19930122
Last Updated on STN: 19970203
Entered Medline: 19921218

AB Three groups of ten calves were each immunised with a total of 400 micrograms pili prepared from three separate strains of *Moraxella bovis* in Alhydrogel-oil adjuvant as two divided, equal doses

THIS PAGE BLANK (USPTO)

10/069799

21 days apart. Groups 1 and 2 each received a monovalent vaccine made from strain 4L and S276R respectively, which belonged to pili serogroup A. Group 3 received vaccine made from pili of strain Maffl, belonging to serogroup F. A further group of ten calves served as non-vaccinated controls. Calves in groups 1 and 2 had developed serogroup A-specific antibody and those in group 3 developed serogroup F-specific antibody, and some evidence of cross-reacting antibody was also detected when measured by an agglutination test using formalin-killed pilated cells of serogroup A strain 4L. Although antibody titres measured against purified pili by ELISA were highest with homologous serogroup antigens, cross-reactive titres to shared epitopes of *M. bovis* pili were also detected by this method. Ocular challenge of the 40 calves with virulent *M. bovis* of serogroup A strain S276R was carried out 14 days after the second vaccine dose. All non-vaccinated calves developed infectious bovine keratoconjunctivitis (IBK). The percentage protection in groups 1 (strain 4L) and 2 (strain S276R) was 60% and 80% respectively (P less than 0.05), with mean lesion scores of 0.7 and 0.3 out of a possible 6.0. The percentage protection of calves in group 3 (strain Maffl) was only 30%, with a mean lesion score of 1.4 compared with 2.2 for non-vaccinated controls. The present findings, together with other evidence indicating that immunity to IBK is serogroup-specific, suggest that inclusion of pili from one representative strain from each of the seven Australian and British serogroups in a polyvalent, subunit vaccine should effectively protect the majority of cattle against IBK caused by most field strains of *M. bovis* encountered in Australia and the United Kingdom.

(FILE 'HCAPLUS, MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH, JICST-EPLUS, JAPIO' ENTERED AT 10:01:50 ON 18 FEB 2004)

L10 45 S (L1 OR L2 OR L3) AND MORAXELLA
L11 0 S L10 NOT L8

FILE 'HOME' ENTERED AT 10:02:32 ON 18 FEB 2004

THIS PAGE BLANK (USPTO)